

From: Mertz, Prema
Sent: Friday, November 08, 2002 4:39 PM
To: STIC-Biotech/ChemLib
Subject: 09/851,595

Please search SEQ ID NO: 5 and 8 with protein databases.

Thanks

Prema Mertz, Ph.D.
Primary Examiner
Art Unit 1646
Crystal Mall 1, Room 10E-01
United States Patent & Trademark Office
(703) 308-4229

RECEIVED
NOV - 8 2002
STIC

Point of Contact
P. Sheppard
Telephone number: (703) 308-4499

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 11/10/02
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2002, 19:32:50 ; Search time 31.4419 Seconds
(without alignments)
4148.213 Million cell updates/sec

Title: US-09-851-595-5
Perfect score: 3307
Sequence: 1 NTHHYRESWYACRYRSGIFG.....GGLSGGGGQPSGLAFASHW 633

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL_21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phase:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_virus:*
17: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3015	91.2	928	4	Q9BYD7
2	2087	63.1	459	11	Q8R301
3	623.5	18.9	1050	5	Q9BN18
4	623.5	18.9	1360	5	Q9ND11
5	608.5	18.4	1300	5	Q9NKB6
6	549	16.6	1012	5	Q9SY16
7	549	16.6	1280	5	Q9SY17
8	459	13.9	701	13	Q9DGC6
9	439	13.3	724	13	Q9VP0
10	432	13.1	779	13	Q918M7
11	428.5	13.0	778	13	Q98TF4
12	418	12.6	658	13	Q9PVN9
13	418	12.6	763	6	Q9BGN4
14	414.5	12.5	693	13	Q9DCC5
15	414	12.5	688	11	Q64183
16	413.5	12.5	831	5	Q94979

17	412	12.5	601	13	Q42500
18	410	12.4	829	5	Q9VE64
19	408	12.3	701	6	Q8SP58
20	408	12.3	764	11	Q9DE97
21	407.5	12.3	831	5	Q8SX01
22	397.5	12.0	793	13	Q91948
23	397	12.0	693	11	Q8R428
24	396	12.0	814	13	Q91949
25	392.5	11.9	662	13	Q9PW16
26	389.5	11.8	739	6	Q9BCE5
27	389.5	11.8	764	6	Q9BC56
28	389.5	11.8	764	6	Q8SP9
29	382.5	11.6	696	13	Q98T85
30	382	11.6	533	4	Q9P1V4
31	381	11.5	699	4	Q15956
32	381	11.5	701	4	Q14751
33	380.5	11.5	696	13	Q9DGF5
34	376	11.4	662	13	Q98T84
35	371.5	11.2	470	4	Q43200
36	366.5	11.1	410	4	Q16225
37	364.5	11.0	929	5	Q18759
38	311	9.4	293	6	Q95MF7
39	291.5	8.8	673	13	Q90WP8
40	284	8.6	335	6	Q46387
41	274	8.3	590	6	Q27986
42	270.5	8.2	307	6	Q95L59
43	259.5	7.8	228	11	Q9ES08
44	255.5	7.7	307	13	Q90XC7
45	181.5	5.5	334	5	Q9VYGO

ALIGNMENTS

RESULT 1

Q9BYD7 PRELIMINARY: PRT: 928 AA.

AC 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE VTS20631 (Fragment).
 GN VTS20631.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Okazaki H., Hayashi A., Kozuma S., Saito T.;
 RT "A member of g-protein coupled receptor family."
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB049405; BAB39854.1; -.
 DR HSSP; P23945; 1XUN.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR003582; LRR-out.
 DR InterPro: IPR003591; LRR-tyr.
 DR Pfam; PF00560; LRR_13.
 DR PRINTS; PR00237; GPCRHHODOPSN.
 DR PRINTS; PR00019; LEURICHRPT.
 DR SMART; SM00370; LRR_9.
 DR SMART; SM00369; LRR_TYR_14.
 FT NON_TER
 SQ SEQUENCE 928 AA: 100487 MW; 4C3364ADEA89C463 CRC64;

Query Match 91.2%; Score 3015; DB 4; Length 928;
 Best local Similarity 99.3%; Pred. No. 1.8e-222;
 Matches 576; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 54 KLEKDFPSQLSLQALDISNAIRSIHPEAFSTLHSLVKLIDLDNDQITPLPLAGLGGLMH 113
 DB 349 EIGADTFPSQLSLQALDISNAIRSIHPEAFSTLHSLVKLIDLDNDQITPLPLAGLGGLMH 408

QY 114 LKLGKGLAQAFAKSDSPKRLILEVPAAYQCCPYGMCASFASQWAEADLHDDSES 173
 DB 409 LKLGKGLAQAFAKSDSPKRLILEVPAAYQCCPYGMCASFASQWAEADLHDDSES 468
 QY 174 SKRPLGLARQAEHNYDDDELQLEMEDSKPHPSVOCSPPTGPPKPCYLLPESWIGIRLA 233
 DB 469 SKRPLGLARQAEHNYDDDELQLEMEDSKPHPSVOCSPPTGPPKPCYLLPESWIGIRLA 528
 QY 234 VMAIVLLSVLCNGVLTVFAGGPAPLPVKFVGAIGANTLTGISCGLASVDALTFEG 293
 DB 529 VMAIVLLSVLCNGVLTVFAGGPAPLPVKFVGAIGANTLTGISCGLASVDALTFEG 588
 QY 294 QFSEYGARWETGLCRATGFLAVLGSEASVLLTLAAVOCSSVSCVRRYKGSPLSGSVR 353
 DB 589 QFSEYGARWETGLCRATGFLAVLGSEASVLLTLAAVOCSSVSCVRRYKGSPLSGSVR 648
 QY 354 AGVIGCLALAGLAALPLASVEYGASPLCLPYAPPEGOPALGFTVALVMNNSFCFLVY 413
 DB 649 AGVIGCLALAGLAALPLASVEYGASPLCLPYAPPEGOPALGFTVALVMNNSFCFLVY 708
 QY 414 AGAYIKLYCDLPRGDFEAVWDCAMVRHVAMLIFFADGLLYCPVAFISFASMLGLFPVTPBA 473
 DB 709 AGAYIKLYCDLPRGDFEAVWDCAMVRHVAMLIFFADGLLYCPVAFISFASMLGLFPVTPBA 768
 QY 474 VKSVLLVPLPACINPLLYLLENPHFRDRLRLPRAGDSGLPAAAGLEKSSCSDST 533
 DB 769 VKSVLLVPLPACINPLLYLLENPHFRDRLRLPRAGDSGLPAAAGLEKSSCSDST 828
 QY 534 QALVAFSDVDLILEASAGRPGLGLEYGFPSTVLTISCOOPGAPRLGSHCVPEEGNHFGN 593
 DB 829 QALVAFSDVDLILEASAGRPGLGLEYGFPSTVLTISCOOPGAPRLGSHCVPEEGNHFGN 888
 QY 594 POPSMDELLLRAGSTPAGGGLSGGGGFGPSGLAFASHV 633
 DB 889 POPSMDELLLRAGSTPAGGGLSGGGGFGPSGLAFASHV 928

RESULT 2
 08R301
 ID 08R301 PRELIMINARY; PRT; 459 AA.
 AC 08R301;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical 47.9 kDa protein (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC026896; AAH26896.1; -.
 KW Hypothetical protein.
 FT NON-TER
 SQ SEQUENCE 459 AA; 47889 MW; F0100BP073E81762 CRC64;

Query Match 63.1%; Score 2087; DB 11; Length 459;
 Best Local Similarity 87.6%; Pred. No. 9.7e-152;
 Matches 402; Conservative 20; Mismatches 37; Indels 0; Gaps 0;

QY 175 KRPLGLARQAEHNYDDDELQLEMEDSKPHPSVOCSPPTGPPKPCYLLPESWIGIRLA 234
 DB 1 KRPLGLARQAEHNYDDDELQLEMEDSKPHPSVOCSPPTGPPKPCYLLPESWIGIRLA 60
 QY 235 WAIIVLLSVLCNGVLTVFAGGPAPLPVKFVGAIGANTLTGISCGLASVDALTFEG 294
 DB 61 WAIIVLLSVLCNGVLTVFAGGPAPLPVKFVGAIGANTLTGISCGLASVDALTFEG 120
 QY 295 FSEYGARWETGLCRATGFLAVLGSEASVLLTLAAVOCSSVSCVRRYKGSPLSGSVR 354
 DB 1 FSEYGARWETGLCRATGFLAVLGSEASVLLTLAAVOCSSVSCVRRYKGSPLSGSVR 354

DB 121 PAEXGARWESGLQCATGFLAVLGSEASVLLTLAAVOCSSVSCVRRYKGSPLSGSVR 180
 QY 355 GVLGCLALAGLAALPLASVEYGASPLCLPYAPPEGOPALGFTVALVMNNSFCFLVY 414
 DB 181 GVLGCLALAGLAALPLASVEYGASPLCLPYAPPEGOPALGFTVALVMNNSFCFLVY 240
 QY 415 GAYIKLYCDLPRGDFEAVWDCAMVRHVAMLIFFADGLLYCPVAFISFASMLGLFPVTPBA 474
 DB 241 GAYIKLYCDLPRGDFEAVWDCAMVRHVAMLIFFADGLLYCPVAFISFASMLGLFPVTPBA 300
 QY 475 KSVLLVPLPACINPLLYLLENPHFRDRLRLPRAGDSGLPAAAGLEKSSCSDSTQ 534
 DB 301 KSVLLVPLPACINPLLYLLENPHFRDRLRLPRAGDSGLPAAAGLEKSSCSDSTQ 360
 QY 535 ALVAFSDVDLILEASAGRPGLGLEYGFPSTVLTISCOOPGAPRLGSHCVPEEGNHFGN 594
 DB 361 ALVAFSDVDLILEASAGRPGLGLEYGFPSTVLTISCOOPGAPRLGSHCVPEEGNHFGN 420
 QY 595 POPSMDELLLRAGSTPAGGGLSGGGGFGPSGLAFASHV 633
 DB 421 POPSMDELLLRAGSTPAGGGLSGGGGFGPSGLAFASHV 459

RESULT 3
 09BN18
 ID 09BN18 PRELIMINARY; PRT; 1050 AA.
 AC 09BN18;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Leucine-rich repeat-containing G protein-coupled receptor 2.
 GN RK OR CG8930.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nishit S., Hsu S.Y., Zell K., Hsueh A.J.;
 RT "Characterization of two fly IGR (leucine-rich repeat-containing G protein-coupled receptor) proteins homologous to vertebrate glycoprotein hormone receptors: constitutively activation of wild type fly IGR1 but not IGR2 in transfected mammalian cells."
 RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF274591; AAK00808.1; -.
 DR FlyBase; FBgn0003255; tk.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR003592; LRR_out.
 DR InterPro; IPR003591; LRR_tyr.
 DR Pfam; PF00001; 7tm_1; 1.
 DR Pfam; PF00560; LRR; 10.
 DR PRINTS; PR00237; GPCR_RHODOPSIN.
 DR SMART; SM00370; LRR; 8.
 DR SMART; SM00369; LRR_tyr; 10.
 DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 1050 AA; 117707 MW; 35D71260A8B4BF99 CRC64;

Query Match 18.9%; Score 623.5; DB 5; Length 1050;
 Best Local Similarity 29.1%; Pred. No. 3.8e-39;
 Matches 174; Conservative 89; Mismatches 234; Indels 101; Gaps 16;

QY 55 LEXDFFSOLSTQALDLSMNAIRSHPEASTLHSLVLTLDNDQLTTLPLAGLGIAMHL 114
 DB 310 LPOAFQGIPIKQLDLDLGNEISYIHKEAFSGFTALDNLGNINIPPELPEGLRALLHD 369
 QY 115 LKLGKGLAQAFAKSDSPKRLILEVPAAYQCCPYGMCASFASQWAEADLHDDSES 173
 DB 370 KTFNNPKLRFPPTFRIOTLITLSTAYHCCAFPLVAMSSQKTSQVQEAIVLPPSDAE 429
 QY 159 -----GQWEADLHDDSESSKRPLGLARQAEHNY 189


```

Db 430 EDTTLMNNSMMNIWPMQHNLSKOLGASHMDPWETA- INFNEQLOTQGTATSYMEY 488
QY 190 DODLDE-----LQEMEDSKPHPSVOCSPPPGPKCEYLFESWGIRLAW 235
Db 489 FEEDHVSGRATGCGTGLFSGMSTEDFQ- GSVQCLPMGPFPLPCADLFDMWTLKCGW 547
QY 236 AIVLLSVLCNGVLLVVFAGGAPRLPPVKFVVGAIAGANTLTGISCGLASVDALTFQGF 295
Db 548 VFLLSLGNGVVFVLLC-SRSKMDVPRFLVGNLAADPFMGITGLAIYDAATLGEF 606
QY 296 SEYGARWETGLGCRATGFLAVLGSSEASVLLTLTAVOCSSVSCVARYKSPSGVRAG 355
Db 607 RMFALIPQMSVLCOLSGFLAVLSSELSV--YTLAVITLERNVAITHAHLNKRSLKQAG 664
QY 356 VLGLL--ALAGLAAPLPLASVGEYASPLCPYPAPREGOPALGFVALVMNSFCFLV 413
Db 665 YIMSVGWVFALIMLMPVGVSDYRKFAVCLPFETTTG-PASLYVISLMEFINGCAFLTL 723
QY 414 AGAYIKLYCDLPRGDFEAVW---DCAMVRHVAWLIFFADGLLYCVAFELSPASMLGLEPVT 470
Db 724 MGCYLMKWAI-RG--SCAMNTNDSRIAKRMALLVFTDFLCWSPIDAFESTITATFGLQLIS 780
QY 471 PEAVKSVLLVLPPLACINPLLYLLENFPERDLRLRPR-----AGDSGPLAYAAAGE 524
Db 781 LEQAKITVFVFLPLNMSCNPFLYAIMTKQFKDCVTLCKHFEESRVYGGGPGRGAVAR 840
QY 525 LEKSSCDSTQALVAFSDVDL---ILEASEAGRPGLFTYGPSTVLLISCOOPGARPLE 579
Db 841 TKRG-----DLPPPLPAAVAHPPGCR-----CLRMLPSEMPNMHKKME 879

RESULT 4
Q9NDI1 PRELIMINARY: PRT: 1360 AA.
AC 09NDI1:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Glycoprotein hormone receptor II.
GN RK OR CG8930.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=CANTON S; TISSUE=WHOLE ANIMAL;
RX MEDLINE=20359836; PubMed=10899142;
RA Erikson K.K., Hauser F., Schiott M., Pedersen K.-M., Soendergaard L.,
RA Grimelikhuisen C.J.P.;
RA "Molecular Cloning, Genomic Organization, Developmental Regulation,
RT and a Knock-Out Mutant of a Novel Leu-Rich Repeats-containing G
RT Protein-Coupled Receptor (Dlcr-2) from Drosophila melanogaster.";
RL Genome Res. 10:924-938(2000).
DR EMBL: AF142343; AAF66608.1;
DR HSP: O57815; ID3Y.
DR FLYBase: FBgn0003255; RK.
DR InterPro: IPR000276; GPCR_Rhodopsin.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR003591; LRR_1yp.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PF00560; LRR_14.
DR PRINTS: PR00237; GPCR_Rhodopsin.
DR SMART: SM00370; LRR_2.
DR SMART: SM00369; LRR_Typ_5.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 1360 AA; 150731 MW; 7DA3515B4F6E12 CR664;
Query Match 18.9%; Score 623.5; DB 5; Length 1360;

```

```

Best Local Similarity 29.1%; Pred. No. 5.2e-39;
Matches 174; Conservative 89; Mismatches 224; Indels 101; Gaps 16;

QY 55 LEQDFSSQLSLOALDLSMNAIRSHPEAFSTLSHLYLDTLDNQLTTLPLAGIGLML 114
Db 524 LPQDAFQGPQLQDLDEGNETSYHKRAFSGFTALDELNGNNIFPLPSGIRALHL 583
QY 115 KIKGNLALSQAFSKDSFPKRLILEYVAYQCC---PYGMCAFFRAS----- 158
Db 584 KTFNNPKLREPPPTFPRIQTLISYAHCCAFPLVAMSSQKTSQVOEAVLFPSDAE 643
QY 159 -----GQWEADHLHDESSKRLGLARQENHY 189
Db 644 EDTTLMNNSMMNIWPMQHNLSKOLGASHMDPWETA- INFNEQLOTQGTATSYMEY 702
QY 190 DODLDE-----LQEMEDSKPHPSVOCSPPPGPKCEYLFESWGIRLAW 235
Db 703 FEEDHVSGRATGCGTGLFSGMSTEDFQ- GSVQCLPMGPFPLPCADLFDMWTLKCGW 761
QY 236 AIVLLSVLCNGVLLVVFAGGAPRLPPVKFVVGAIAGANTLTGISCGLASVDALTFQGF 295
Db 762 VFLLSLGNGVVFVLLC-SRSKMDVPRFLVGNLAADPFMGITGLAIYDAATLGEF 820
QY 296 SEYGARWETGLGCRATGFLAVLGSSEASVLLTLTAVOCSSVSCVARYKSPSGVRAG 355
Db 821 RMFALIPQMSVLCOLSGFLAVLSSELSV--YTLAVITLERNVAITHAHLNKRSLKQAG 878
QY 356 VLGLL--ALAGLAAPLPLASVGEYASPLCPYPAPREGOPALGFVALVMNSFCFLV 413
Db 879 YIMSVGWVFALIMLMPVGVSDYRKFAVCLPFETTTG-PASLYVISLMEFINGCAFLTL 937
QY 414 AGAYIKLYCDLPRGDFEAVW---DCAMVRHVAWLIFFADGLLYCVAFELSPASMLGLEPVT 470
Db 938 MGCYLMKWAI-RG--SCAMNTNDSRIAKRMALLVFTDFLCWSPIDAFESTITATFGLQLIS 994
QY 471 PEAVKSVLLVLPPLACINPLLYLLENFPERDLRLRPR-----AGDSGPLAYAAAGE 524
Db 995 LEQAKITVFVFLPLNMSCNPFLYAIMTKQFKDCVTLCKHFEESRVYGGGPGRGAVAR 1054
QY 525 LEKSSCDSTQALVAFSDVDL---ILEASEAGRPGLFTYGPSTVLLISCOOPGARPLE 579
Db 1055 TKRG-----DLPPPLPAAVAHPPGCR-----CLRMLPSEMPNMHKKME 1093

RESULT 5
Q9NRD6 PRELIMINARY: PRT: 1300 AA.
AC 09NRD6; Q9YJ03;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 144.0 kDa protein (Rk gene product).
GN RK OR CG8930.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=BERKELEY;
RX MEDLINE=99403001; PubMed=10471707;
RA Ashburner M., Mista S., Noote J., Lewis S.E., Blazek R., Davis T.,
RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
RA Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitelaw K.,
RA Gelinkner S., Rubin G.M.;
RA "An exploration of the sequence of a 2.9-Mb region of the genome of
RT Drosophila melanogaster: the Adh region.";
RL Genetics 153:179-219(1999).
KW
SQ SEQUENCE FROM N.A.
RP STRAIN=BERKELEY;

```


DR	PROSITE	PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
DR	PROSITE	PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR	PROSITE	PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR	RECEPTOR	
DR	SEQUENCE	1280 AA; 141700 MW; 0ABDECCDD880BA CRC64;
QY	55	LEQDTFSQLSLQALDLSMNAINSHIEAFSTLSHYKLDITDNOITLPTPLAGLGGMLH 114
DB	369	IPADAFOSLSHLDTLSLSNNITREIDSQAAPQTSLSQYLDLSNNSFPVPTPLGLOMLKTI 428
QY	115	KLKGNGLMSAQFSNDSPPKRIILEVPAAYCCPGKCAQSFKASGMEADYHLHDEESS 174
DB	429	KTYDEQLEDEPPPESELPSTIEITATPYHCCEYIELAEFLKS-----LADRPNI 479
QY	175	KRPGLLLARQAEHHYDDDLDELQLEMEDS-----KPRHSVC 211
DB	480	SETTYWAGSVDPDYNNMTFIDINSESWIDSIFGEGSLSIGSPYLSGNSYSLVPH-NISC 538
QY	212	SPTPEPRKCEYLFESMGIRLAWAYILLVSLNGVILLVFPAGGPAPLPVKEFVGAIA 271
DB	539	RKRPEPRPCMDLFGSNPLRGVPLVLLAIGALTYIPIIYSHRKMDVP-RELICNIA 597
QY	272	GANTLTGTGSCGLASVDALFRGPQSEYGARWETGLGCRATGFLAVLGSSEASVLLTAAV 331
DB	596	FADPLFLGYIYGLAGVDSTLGVFRKRGARQWLSAGCRLAGFLAVFSEFSIYTLSTVTL 657
QY	332	QCSVSVCVRATGKSPSLGSRAGVIGCLA--LAGLAAALPLASVGEYGAAPCLPYAPP 389
DB	658	ERRVYAIKHAHLHEKRMK--PHATIVMCEGFIESTVAIVLPLVNVSHHYKVPCLPF--D 713
QY	390	EQOPALGFTVALVMANSEFCVLVAGATIKLYCDLPGRDEEAVWDC--AMVSHVAMLV 446
DB	714	VDTYAKVYVGSILLNLNLAIVITIMACYASITYLAI-QGSH--AMNCDSVVARMSILVF 770
QY	447	ADGLLYCPVAPLIFSRASMTGLPPVPEAVKSVLLVYLPPLACTNLNLLPLNPHRDLRR 506
DB	771	TDFACWAPIAFPSLTAAGFLRLISDGAKVLTIVPLPLNSCANPFLYITLTKQFKDKCT 830
QY	507	L 507
DB	831	I 831

Query Match 16.6%; Score 549; DB 5; Length 1280;
 Best Local Similarity 28.5%; Pred. No.2.5e-33;
 Matches 137; Conservative 78; Mismatches 220; Indels 46; Gaps

RESULT 8
 Q9DGC6 PRELIMINARY; PRT; 701 AA.

AC Q9DGC6: 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Gonadotropin receptor 1.
 GN TGT8-RI.
 OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorphi; Perciformes; Labroidae;
 OC Cichlidae; Oreochromis.
 OX NCBI_TaxID=8128;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Oba Y., Hirai T., Yoshiura Y., Yao Z., Nagahama Y.;
 RL "Tilapia gonadotropin receptor 1";
 RL Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AB041762; BAB16106.1; -;
 DR HSSP; P23945; 1XUN.
 DR Interpro; IPR000276; GPCR_Rhodpsn.
 DR Interpro; IPR001611; LRR-
 DR Pfam; PF00560; LRR; 4.
 DR Pfam; PF00560; LRR; 4.


```

Db 670 LVLVFLPLNSCANPFLVAFITKAFRGDIFLLSKVG 705

RESULT 12
Q9PVN9 PRELIMINARY; PRT; 658 AA.
ID Q9PVN9
AC Q9PVN9;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
DE Gonadotropin receptor I.
GN SGT8-RI.
OS Oncothynechus rhodurus (amago).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncothynechus.
OX NCBI_TaxID=41164;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20025349; PubMed=10558873;
RA Oba Y., Hirai T., Yoshura Y., Yoshikuni M., Kawachi H., Nagahama Y.;
RT "The duality of fish gonadotropin receptors: cloning and functional
RT characterization of a second gonadotropin receptor cDNA expressed in
RL the ovary and testis of amago salmon (Oncothynechus rhodurus).";
RL Biochem. Biophys. Res. Commun. 285:366-371(1999).
DR EMBL: AB030012; BAA86898.1; -.
DR HSSP: P22888; 1LUT.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR001611; LRR.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PF00560; LRR_2.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 658 AA; 72991 MW; 853A30BEA845480 CRC64;

Query Match 12.6%; Score 418; DB 13; Length 658;
Best Local Similarity 24.7%; Pred. No. 1.2e-23;
Matches 139; Conservative 84; Mismatches 238; Indels 102; Gaps 10;

QY 16 SGIPGTHASVERSGSLPAHPASIALAASNTTA--SGLENDTFEQLSS----- 65
DB 98 ANPRLTEITTKSKHVIHQAFLGPKLSHLTCTGTGLVLPNFSRTHSAAMTFLD 157
QY 66 -----LQALDLSMNAIRSIHPEAF--STLHSLVK----- 92
DB 158 LQDNVHVIIPNAFLGLTNTIDELRLTKNGISEVESHAENGTRIKHLIMGNQLSHM 217
QY 93 -----LDLTDNQLTTPLAGLGLMLKLGKGNLALSOAFSKDSFPKRLILEV 139
DB 218 HNSSEFGAEGPGLDLSRLALSSLPESVLGEVHLASAVSFSLRPLPLSLFTKRLQANL 277
QY 140 PYAVQCC-----PYGCASFEEKASGQWEAEADLHLDDESSKRPGLGLAQEENYD 190
DB 278 TVPSHCCAFHKHQNRFRFTMSACFKGAQ----- 307
QY 191 QDDDELQLEMEDESKRHPVSQCCSPPPGPKCEYLPESWGIRLAWAIVLLSVCNGVLV 250
DB 308 ---NNHAFEDFLCMTWSTVACSPAPDAFNPCEDIMGSAFLPILMIITISVLLGLTIVL 364
QY 251 TVPAGGAPRLPYKFFVGAAGANTLTGICGLASVDALTFGQFSEYGAERWETGLGGR 310
DB 365 -VLLGSRAKKTVPRLMLCHLSPADLCMGITLVATVDVFRGLYVNIATISWGTAGCDI 423
QY 311 TGFVLAVGSEASVLLTLAIVQCSVSVAYKSPSLGSRAGVLCCLALAGLAALP 370
DB 424 AGFFTVFASLSMFTLITATLERYHTTHRLDRKLRLRACAVMTGAFSCLAALLP 483
QY 371 LASVGEYGAFLCLPYAPPGOPALAGFTYALVWMSFCELVAGAVIKLYCDLPRGDF 429
DB 484 TVGVSSYSKVSICLP-MVESLPSQV-FVMFLLLNVVAFLVCVCYLSTILSVRNSSSP 541

```

```

QY 430 EAVWDCAWVRVAVLITADGLLYCPVAFLSRSMIGLFPVPEAVKSVLLVLPAPACIN 489
DB 542 PASAETCMAQFMALITFDLCMAPISEFFALSAALIKLPLTVSDSKLLVLYFPINSCAN 601
QY 490 PLIYLFPNHPRDDLRRLPRAG 512
DB 602 PFLYGLCTRTFRDRDFLLAARYG 624

RESULT 13
Q9BGN4 PRELIMINARY; PRT; 763 AA.
ID Q9BGN4
AC Q9BGN4;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
DE Thyrotropin receptor.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RX Nguyen L.Q., Karamanoglu Arseven O., Gerber H., Stein B.S.,
RA Jameson J.L., Kopp P.;
RT "Cloning of the cat thyrotropin receptor and evidence against an
RT autoimmune etiology of feline hyperthyroidism.";
RT Submitted (DEC-1999) to the EMBL/GenBank/DBS databases.
DR EMBL: AF218264; AAK0133.1; -.
DR HSSP: P16473; 1XDM.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR001611; LRR.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PF00560; LRR_1.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 763 AA; 86588 MW; 937F60A140487D60 CRC64;

Query Match 12.6%; Score 418; DB 6; Length 763;
Best Local Similarity 24.4%; Pred. No. 1.4e-23;
Matches 144; Conservative 108; Mismatches 254; Indels 84; Gaps 15;

QY 33 SLPAHPASIALAASNTTASGKLENDTFEQL-----SLQALDLSMNA-IRSIHPEAF 84
DB 166 SIPAN-----AFQSLCNETLTLKLYKNGFTSIGAHAFNGTKLDVAIVLLKNNKYLITADDAF 221
QY 85 STLHS-LVKLDLTDNQLTTPLAGLGLMLKLGKGNLALSOAFSKDSFPKRLILEVYAY 143
DB 222 GGYYSGPTLLDVSYSTVALPSKGLHKLRLARNWTLLKPLTLISFLHUTRADLSYPS 281
QY 144 OCCPY-----GMCASF-----KASGWEAE-----DLHLDDESS 174
DB 282 HCCAFKNQKKIRKILSEFMCNDSSIRSLRQKSVNALNGFDEYEBEYLDSDSHAGYDNS 341
QY 175 KRPLGLAQEENHY-----DQDL-----DELOLEMEDESKR-----HPSVQC 211
DB 342 K-----FQDTRNSHYVVFEEDEDEILGFGQELKNPQETLQAFDSHHDYTVCGNENMVC 397
QY 212 SPYPPGPKCEYLPESWGIRLAWAIVLLSVCNGVLVTFVAGGAPRLPYKFFVGAIA 271
DB 398 TPKSDEFNCEIDIMGYKFLVYVWFSLLALGNVFLVLLITLSHYKLTVP-REFMCNLA 456
QY 272 GANTLTGICGLASVDALTFGQFSEYGAERWETGLGCRATGFLAVGSEASVLLTLTAAY 331
DB 457 FADFCMGWTLILLASVDLTHSEYVNHAIQWGTGPGCNAAGFTTVFASLSVTVTLVITL 516
QY 332 QCSVSVSCVAVKSPSLGSRAGVLCCLALAGLAALPLASVGEYGASPLCLPYAPPEG 391
DB 517 ERWYATFAMRLDRKMRRLRHAYAIWVGWVCCFLALLPLLVGISYAKVSICLPM--DTE 574
QY 392 QPALGFTYALVWMSFCELVVAGAVIKLYCDLPRGDF-AVWDCAWVRVAVLITADGL 450

```

```

Db 575 TPLALATIIIVLLINIAVAFIVCGYKITYITVANPCYNTGDKDKRIAKRAVALIFTEDEM 634
OY 451 LYCPVAFELSEASMLGEPVTPPEAVKSVLLVLPACLNPLIYLLFNPHFRDDLRLRAPP 510
Db 635 CMADISFVALSALMNKPLITVTNSKILLVLEFPYNSCANPFLYALFTFTFPFIDLLSK 694
OY 511 AGDSGSPLAYAAG-----ELEKSGDSGQALVAFSPVDILBAS 549
Db 695 FGICKRQAOAYRGORVSPKSTGIQVQKVTNRNROSILPRMODDYELLNS 744

RESULT 14
ID 09DGC5 PRELIMINARY: PRT: 693 AA.
AC 09DGC5;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE Gonadotropin receptor II.
GN TGNH-RII.
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphi; Acanthopterygii; Perciformes; Labroidae;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RA Oba Y., Hirai T., Yoshitura Y., Yao Z., Nagahama Y.:
RT "Tilapia gonadotropin receptor II."
RL Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AB041763; BAB16107.1; -.
DR HSSP: P22888; ILUT.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR001611; LRR.
DR Pfam: PF000001; 7tm_1; 1.
DR Pfam: PF00560; LRR; 2.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR PROSITE: PS00237; G_PROTEIN_RECPEP_FL_1; UNKNOWN_1.
DR PROSITE: PS0262; G_PROTEIN_RECPEP_FL_2; 1.
DR Receptor.
KW SEQUENCE 693 AA; 77856 MW; C3BBDP8EFCF8988 CRC64;

Query Match 12.5%; Score 414.5; DB 13; Length 693;
Best Local Similarity 28.0%; Pred. No. 2.3e-23;
Matches 134; Conservative 79; Mismatches 204; Indels 61; Gaps 14;

OY 61 SOLSSIQALDLSNNAIRSHPEAFSTLHSLVKLDLTDNQLTTPPLACLGSLMHLKLGNL 120
Db 200 TKLINTVLNRD-NM-YLRNIQEDAFEGATGPTLLDVSTALRSLP--PGLRHVKF--L 251
OY 121 ALSQASKSPFRLRL-----EVPYAYOCPCYGMCAFPFKASGOWEADLHIDDESS 174
Db 252 KASHAVALSLPLLESIAELLELELYPSHC-----AFHWRRK-----QRESA 296
OY 175 KRPLGLLAROENHYDODDELQI-----EME-DSKPHPSVOCSPRPGRPKCEYL 224
Db 297 LKNLWLFCDLMTNEIPTADTSLINDINFOYDPLEDFDCSNPFVKSPRDAFNPCEDL 356
OY 225 FESEWGRILAWAIVLVSLVLCNGVLVLFVAGCAPRLPVKFGVIGALGAMTLGSGGL 284
Db 357 LGFSFRCCLTWITMVAVAGN-LAVLVILLIGHKRLVSRFLKCNLAFADLCMGLYILI 415
OY 285 ASVDALTFQGFSEAGRWETGLCRATGFLAVLGSSEASVLLTLTAAVQCSVSVCAVAG 344
Db 416 AFMDYHSHHHYNNHATPWQGGPGGGLGFLVPSSELSVTLTVISLEPMHTTNAMHVN 475
OY 345 KSPSLGSVRAVLGCLALAGLAALPLASVGEYASPLCLPYARPEQPAALG---FTVA 401
Db 476 KRLRMHHVTAMVGMGAFSLIVALLPLVGVSSYKSVISCLPM---DIDTLGAQVYVVA 530
OY 402 LVMAANSFCFLVAVAGYIKIKIDL-----PRQFVAVWDCAMVRHYVAMLIFADGLLYCP 454

```

```

Db 531 VLIINVAFLVWCYCYIGIYISVHNPEHSTRGRDK-----IAKRMVLIPTDPLCAP 584
OY 455 VAFLEASMLGEPVTPPEAVKSVLLVLPACLNPLIYLLFNPHFRDDLRLRAPP 512
Db 585 ISFEAISALRNPLITVSHSKILLIFPINSCLNPFLYITFTFRFRKDYCLLSRCG 642

RESULT 15
ID 064183 PRELIMINARY: PRT: 688 AA.
AC 064183;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE Follicle-stimulating hormone receptor.
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92149579; Pubmed=1738373;
RA Heckert L.V., Daley I.J., Griswold M.D.;
RT "Structural organization of the follicle-stimulating hormone receptor
RT gene."
RL Mol. Endocrinol. 6:70-80(1992).
DR EMBL: S81198; AAB21415.2; -.
DR EMBL: S81117; AAB21415.2; JOINED.
DR EMBL: S81119; AAB21415.2; JOINED.
DR EMBL: S81121; AAB21415.2; JOINED.
DR EMBL: S81171; AAB21415.2; JOINED.
DR EMBL: S81174; AAB21415.2; JOINED.
DR EMBL: S81178; AAB21415.2; JOINED.
DR EMBL: S81183; AAB21415.2; JOINED.
DR EMBL: S81185; AAB21415.2; JOINED.
DR EMBL: S81194; AAB21415.2; JOINED.
DR HSSP: P23945; 1XUN.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR00372; LRR_Nterm.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PF00560; LRR; 3.
DR Pfam: PF01462; LRRNT; 1.
DR SMART: SM00013; LRRNT; 1.
DR PROSITE: PS00237; G_PROTEIN_RECPEP_FL_1; UNKNOWN_1.
DR PROSITE: PS0262; G_PROTEIN_RECPEP_FL_2; 1.
DR Receptor.
KW SEQUENCE 688 AA; 77341 MW; 441F0D9E7D01DF18 CRC64;

Query Match 12.5%; Score 414; DB 11; Length 688;
Best Local Similarity 24.8%; Pred. No. 2.5e-23;
Matches 165; Conservative 86; Mismatches 253; Indels 160; Gaps 24;

OY 16 SGIGSTHNASVERSGQ-----SLPAHPASIALAASNTTASGLKEXDFQSLSL 66
Db 92 SNLPKLEIRKRNANLLYNPEAFQNLN-----SLRVLISNT-----GIKHLPAVHKIOSL 144
OY 67 Q-----ALDSW-----NAIRSHPEAF-----STLHS 89
Db 145 QKVLLDIQDININIIIVARNSEMGISFEMLSKNGIEEIHNCANFGSTQDELMSNNNIEE 204
OY 90 L-----VKLDLTDNQLTTPPLAGLGLMHLKGNLALASQASKSPFRLRL 138
Db 205 LPNDVFGASPVLLDSIRTVSHSLPNHGLENLKRLARSRVYRKKLPNDLKFVTLMEAS 264
OY 139 VPYAYOCPCYGMCAFPFKASGOWEADLH-----LDDDES 173
Db 265 LTFPSGCCAF-----ANLKRQISELHPICNKSILRQDIDMTQIGDQVSLIDEPS 316
OY 174 SKRPLGLLAROENHYDODDELQLEMDSKPHPSVOCSPRPGRPKCEYLFESEWGRILA 233
Db 317 ---YKGSDDMYNMEFDYLCN---EYVD-----VTCSPKPRDAFNPCEDMIGINILRYL 363

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2002, 19:31:55 ; Search time 9.24763 Seconds
(without alignments)
2839.055 Million cell updates/sec

Title: US-09-851-595-5

Perfect score: 3307
Sequence: 1 NTTHRESWYACRYSRGPIC.....GGLSGGGGRQPSGLAFASHV 633

Scoring table: BLASTSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SWISSPROT_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3011	91.0	828	1	LG86_HUMAN
2	1275	38.6	907	1	LG85_HUMAN
3	1226	37.1	907	1	LG85_MOUSE
4	967	29.2	951	1	LG84_RAT
5	957	28.9	951	1	LG84_HUMAN
6	440	13.3	692	1	FSHR_MOUSE
7	428.5	13.0	694	1	FSHR_HORSE
8	424.5	12.8	695	1	FSHR_BOVIN
9	422	12.8	695	1	FSHR_PIG
10	419	12.7	696	1	FSHR_SHEEP
11	418.5	12.7	695	1	FSHR_RAT
12	417	12.6	692	1	FSHR_RAT
13	414.5	12.5	695	1	FSHR_HUMAN
14	414.5	12.5	695	1	FSHR_MACFA
15	414.5	12.5	764	1	TSHR_CANFA
16	414.5	12.5	764	1	TSHR_SHEEP
17	412.5	12.5	700	1	LSHR_MOUSE
18	409	12.4	763	1	TSHR_BOVIN
19	408	12.3	764	1	TSHR_MOUSE
20	407.5	12.3	700	1	LSHR_RAT
21	404	12.2	701	1	LSHR_BOVIN
22	402	12.2	687	1	FSHR_EQUAS
23	398.5	12.1	764	1	TSHR_HUMAN
24	381.5	11.5	764	1	TSHR_RAT
25	381	11.5	699	1	FSHR_HUMAN
26	377.5	11.4	693	1	FSHR_CHICK
27	377	11.4	676	1	LSHR_CALJA
28	363	11.0	538	1	LSHR_SHEEP
29	357	10.8	366	1	LSHR_CHICK
30	309	9.3	754	1	LG88_HUMAN
31	308.5	9.3	925	1	GLHR_ANCEL
32	306.5	9.3	737	1	LG88_MOUSE
33	299	9.0	1115	1	GPCR_LYMST

34	269.5	8.1	757	1	IGR7_HUMAN
35	149.5	4.5	536	1	CBP8_HUMAN
36	138	4.2	646	1	FLR1_HUMAN
37	137.5	4.2	355	1	OPSB_GECGE
38	135.5	4.1	370	1	GALT_RAT
39	134.5	4.1	370	1	GALS_RAT
40	134	4.1	354	1	OPSD_GALML
41	134	4.1	371	1	GALS_MOUSE
42	133.5	4.0	370	1	GALT_MOUSE
43	132	4.0	377	1	APJ_MOUSE
44	132	4.0	377	1	APJ_RAT
45	131	4.0	350	1	MLIA_HUMAN

ALIGNMENTS

RESULT 1
ID LG86_HUMAN STANDARD: PRT; 828 AA.

AC Q9H8X8; Q96K69; 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DE 15-JUN-2002 (Rel. 41, Last annotation update)

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. MEDLINE=20388592; PubMed=10935549;

RA Hsu S.-Y., Kudo M., Chen T., Nakabayashi K., Bhalla A., Tanase T., Nomura Y., Toshiya S., Komai F., Hara R., Takeuchi K., Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y., Oshima A.;

RA "NEBO human cDNA sequencing project." Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

RT FUNCTION: Orphan receptor.

RL SUBCELLULAR LOCATION: Integral membrane protein.

CC SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC SIMILARITY: CONTRAINS 11 LEUCINE-RICH REPEATS (LRR).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: AF190501; AAG17168.1; -

CC EMBL: AB083616; BAB89329.1; -

CC EMBL: AK027377; BAB55071.1; ALR_INIT.

CC MIM: 606653; -

CC Interpro: IPR000276; GPCR_Rhodopsn.

```

DR InterPro: IPR001611; LRR.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR003591; LRR_tyr.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PF00560; LRR; 8.
DR PRINTS: PRO1561; EDG8RECEPTOR.
DR PRINTS: PRO0237; GPCR_HODOPSN.
DR PRINTS: PRO0019; LEURICHRPT.
DR SMART: SM00370; LRR; 5.
DR SMART: SM00369; LRR_TYP; 10.
DR PROSITE: PS00237; G_PROTEIN_RECPT_F1_1; FALSE_NEG.
DR PROSITE: PS00262; G_PROTEIN_RECPT_F1_2; FALSE_NEG.
DR G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat;
KW Leucine-rich repeat.
FT DOMAIN 1 428
FT TRANSSEM 429 449
FT DOMAIN 450 459
FT TRANSSEM 460 480
FT DOMAIN 481 505
FT TRANSSEM 506 526
FT DOMAIN 527 548
FT TRANSSEM 549 569
FT DOMAIN 570 588
FT TRANSSEM 589 609
FT DOMAIN 610 635
FT TRANSSEM 636 656
FT DOMAIN 657 670
FT TRANSSEM 671 691
FT DOMAIN 692 804
FT TRANSSEM 813 816
FT DOMAIN 822 45
FT REPEAT 46 69
FT REPEAT 71 93
FT REPEAT 94 117
FT REPEAT 118 140
FT REPEAT 142 164
FT REPEAT 168 211
FT REPEAT 212 236
FT REPEAT 238 257
FT REPEAT 258 281
FT REPEAT 283 305
FT DISULFID 503 578
FT CARBOHYD 15 15
FT CARBOHYD 34 34
FT CONFLICT 406 410
FT CONFLICT 628 628
FT CONFLICT 824 824
SQ SEQUENCE 828 AA; 89301 MW; 1B5971445AAD8B4 CRC64;

```

```

Query Match 91.0%; Score 3011; DB 1; Length 828;
Best Local Similarity 99.1%; Pred. No. 2.5e-214;
Matches 575; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 54 KLEXTFQSLSQALDLSWNAIRSIHEAFSTLSYKLDITDQTLTLPAGLGGLMH 113
DB 249 ELGADTFQSLSQALDLSWNAIRSIHEAFSTLSYKLDITDQTLTLPAGLGGLMH 308
QY 114 LKLGKGNLALSOAFKSDSPKRLILEVPAVOCPCPGKMAFFKASGQWEADLHDDDEES 173
DB 309 LKLGKGNLALSOAFKSDSPKRLILEVPAVOCPCPGKMAFFKASGQWEADLHDDDEES 368
QY 174 SKRPLGLARAEHNYHDDDELDELEMEDSKPHPSVOCSPTPGPKPEVLFESGIRLA 233
DB 369 SKRPLGLARAEHNYHDDDELDELEMEDSKPHPSVOCSPTPGPKPEVLFESGIRLA 428
QY 234 VMAIVLVLVLCNGVLVLTFFAGGPAIPDVVGVGATGANTLTIGISGGLASVDALTFG 293
DB 429 VMAIVLVLVLCNGVLVLTFFAGGPAIPDVVGVGATGANTLTIGISGGLASVDALTFG 488
QY 294 QFSYEGARWETGLGCRATGFLAVGSEASVLLTLAAVOCVSVSCVAVGSKSPSLGSVR 353
DB 489 QFSYEGARWETGLGCRATGFLAVGSEASVLLTLAAVOCVSVSCVAVGSKSPSLGSVR 548

```

```

QY 354 AGVLGLALAGLAAALPLASVGEYASPLCLPYAPPEGOPALGFTVALVMNSCEFLV 413
DB 549 AGVLGLALAGLAAALPLASVGEYASPLCLCLPYAPPEGOPALGFTVALVMNSCEFLV 608
QY 414 AGATIKLYCDLPGRDFEAVWDCAVRYHVAWMLIFADGLLYCPVAFSLFASMLGLFPVTEA 473
DB 609 AGATIKLYCDLPGRDFEAVWDCAVRYHVAWMLIFADGLLYCPVAFSLFASMLGLFPVTEA 668
QY 474 VKSVLVLPACINPLLYLLFNPHFRDDLRRLPRRAGDSGPLAYAAAGLEKSCDST 533
DB 669 VKSVLVLPACINPLLYLLFNPHFRDDLRRLPRRAGDSGPLAYAAAGLEKSCDST 728
QY 534 QALVAFSDVDLILEASEGRPPLETETYPSPVTLISCOOPGAPRLGSHCVPEEGNHFEN 593
DB 729 QALVAFSDVDLILEASEGRPPLETETYPSPVTLISCOOPGAPRLGSHCVPEEGNHFEN 788
QY 594 POPSMDEGLLRARGSPAGGSLGGGFGFOPSGLAFAFASHV 633
DB 789 POPSMDEGLLRARGSPAGGSLGGGFGFOPSGLAFAFASHV 828

```

RESULT 2

IGRS_HUMAN STANDARD; PRT; 907 AA.

```

AC 075473; Q9UP75;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Leucine-rich repeat-containing G protein-coupled receptor 5 precursor
DE (Orphan G protein-coupled receptor HG38) (G protein-coupled receptor
49).
DE GPR49 OR IGR5 OR GPR67.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98308104; PubMed=9642114;
RA McDonald T., Wang R., Bailey W., Xie G., Chen F., Caskey C.T.,
RA Liu Q.;
RT "Identification and cloning of an orphan G protein-coupled receptor of
the glycoprotein hormone receptor subfamily."
RL Biochem. Biophys. Res. Commun. 247:266-270(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Placenta;
RX MEDLINE=99065210; PubMed=9849958;
RA Hsu S.Y., Liang S.-G., Hsueh A.J.W.;
RT "Characterization of two IGR genes homologous to gonadotropin and
thyrotropin receptors with extracellular leucine-rich repeats and a G
protein-coupled, seven-transmembrane region."
RL Mol. Endocrinol. 12:1830-1845(1998).
RN [3]
RP FUNCTION: Orphan receptor. It may be an important receptor for
signals controlling growth and differentiation of specific
embryonic tissues (by similarity).
RN [4]
RP SUBCELLULAR LOCATION: Integral membrane protein.
RN [5]
RP TISSUE SPECIFICITY: Expressed in skeletal muscle, placenta, spinal
cord, and various region of brain.
RN [6]
RP SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
RN [7]
RP SIMILARITY: CONTAINS 17 LEUCINE-RICH REPEATS (LRR).

```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: AF062006; AAC28019.1; -
 EMBL: AF061444; AAC77911.1; -
 Genew: HGNC:4504; GPR49.

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL; AF110818; AAD14684.1; -
 DR HSSP; P23945; 1XUN.
 DR MGD; MG1:1341817; Gp49.
 DR InterPro: IPR000276; GPCR_Rhodopsin.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Nterm.
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR003591; LRR_typ.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 15.
 DR Pfam: PF01462; LRRNT; 1.
 DR PRINTS; PR00237; GPCR_RHODOPSIN.
 DR PRINTS; PR00019; LEURICHRPT.
 DR SMART; SM00370; LRR; 3.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_TYP; 8.
 DR PROSITE; PS00237; G-PROTEIN_RECEP_FL_1; FALSE_NEG.
 DR PROSITE; PS00262; G-PROTEIN_RECEP_FL_2; 1.
 DR G-protein coupled receptor; Signal; Transmembrane; Glycoprotein;
 KW Repeat; Leucine-rich repeat.
 FT CHAIN 1 22 907
 FT SIGNAL 1 21
 FT DOMAIN 22 561
 FT TRANSMEM 562 582
 FT DOMAIN 583 593
 FT TRANSMEM 594 614
 FT DOMAIN 615 638
 FT TRANSMEM 639 659
 FT DOMAIN 660 682
 FT TRANSMEM 683 703
 FT DOMAIN 704 723
 FT TRANSMEM 724 744
 FT DOMAIN 745 767
 FT TRANSMEM 768 788
 FT DOMAIN 789 802
 FT TRANSMEM 803 823
 FT DOMAIN 824 907
 FT REPEAT 89 112
 FT REPEAT 113 136
 FT REPEAT 137 160
 FT REPEAT 162 184
 FT REPEAT 186 208
 FT REPEAT 209 232
 FT REPEAT 233 256
 FT REPEAT 257 279
 FT REPEAT 281 303
 FT REPEAT 304 327
 FT REPEAT 328 350
 FT REPEAT 351 375
 FT REPEAT 377 396
 FT REPEAT 397 420
 FT REPEAT 422 444
 FT REPEAT 444 465
 FT CARBOHYD 63 63
 FT CARBOHYD 77 77
 FT CARBOHYD 208 208
 FT CARBOHYD 792 792
 SO SEQUENCE 907 AA; 99681 MW; 553167C6C0AAE253 CRC64;

Query Match 37.1%; Score 1226; DB 1; Length 907;
 Best Local Similarity 46.3%; Pred. No. 1; le-82;
 Matches 244; Conservative 102; Mismatches 163; Indels 18; Gaps 8;
 54 KLEADTSQSLQALDLSMAINSIHPEAFSTLHSLVKIDLTNDQTLTFLAGLGLMH 113

Db 388 ELKSTFOOLFRLRLNLAWNKIALIHHNAFSTPLIKDLSSLLSPFTYGLHGLH 447
 Qy 114 LKLNGLNSQAFSDSPKRIIEVPYACCPKMGKASFFKASGQWAEADLHDEES 173
 Db 448 LKLNGLNSQAFSDSPKRIIEVPYACCPKMGKASFFKASGQWAEADLHDEES 173
 Qy 174 SKRPLGLLARGAENHYDDDELQLEM-EDSKPHPSVQCSPTPPGPKCEYLFEESWGIRL 232
 Db 507 HKKDALGFQVQDE---RDLEDPLDPEEDLNALHSVQCSPPGPKCEYLFEESWGIRL 562
 Qy 233 AVMAVLLSVLCNGVLTVPAGGAPAPLPVKFVVGAIAGANTLTGISGLLAVDALTF 292
 Db 563 GVMVAVLLSVLCNGVLTVPAGGAPAPLPVKFVVGAIAGANTLTGISGLLAVDALTF 292
 Qy 293 GQFSEYGAEMETGLCRATGFLAVIGSEASVLLTLTAAVQCSVSCVRAKGPISLGSV 352
 Db 622 GRFAGHGMWEDGICQVIGFLSIFASFSFLTLTALERGSVKSKFEVKAFLFSL 681
 Qy 353 RAGVGLCLALAGLALPLASVGEYASPLCLPYAPPEGAPALGFVVALVMNSFCFLV 412
 Db 682 RAIVLLCVLLATLTATPILGGSKYKNSPCLPL-PPGEPSTTGVMVALVLLNSLCLFLI 739
 Qy 413 VAGAYIKLYCDLPRGDEAVWDCAVRHVAMLTFRADGLTCVPAFLSFAISLGFVPTPE 472
 Db 740 MTIATVTLKYSLEGELENMDCSMVKHIALDLFANCLTCVPAFLSFAISLGFVPTPE 799
 Qy 473 AVKSVLLVPLPACNPLLYLLENPHFRDLRLRLRRA-----GDSGPLAYAAAGLEK 527
 Db 800 VIKRILLVPLPACNPLLYLLENPHFRDLRLRLRRA-----GDSGPLAYAAAGLEK 527
 Qy 528 SSCDSPTALAFSPVDLI--LEASAGRP--PGLETGFPFVLLISC 570
 Db 860 RSCDSPTALAFSPVDLI--LEASAGRP--PGLETGFPFVLLISC 570
 RESULT 4
 LGR4_RAT STANDARD; PRT; 951 AA.
 ID LGR4_RAT
 AC 0922H4;
 DT 15-JUN-2002 (rel. 41, Created)
 DT 15-JUN-2002 (rel. 41, Last sequence update)
 DT 15-JUN-2002 (rel. 41, Last annotation update)
 DE Leucine-rich repeat-containing G protein-coupled receptor 4 precursor.
 GN GPR48 OR LGR4.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RX MEDLINE=99065210; PubMed=9849958;
 RA Hsu S.Y., Liang S.-g., Hsueh A.J.W.;
 RT "Characterization of two LGR genes homologous to gonadotropin and
 RT thyrotropin receptors with extracellular leucine-rich repeats and a G
 RT protein-coupled, seven-transmembrane region.";
 RL MOL. Endocrinol. 12:1830-1845(1998).
 CC -1- FUNCTION: Orphan receptor.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -1- SIMILARITY: CONTAINS 15 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

EMBL; AF061443; AAC77910.1; -
 InterPro: IPR000276; GPCR_Rhodopsin.

DR InterPro: IPR001611; LRR.
 DR InterPro: IPR003572; LRR_Nterm.
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR003591; LRR_Typ.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 15.
 DR Pfam: PF01462; LRRNT; 1.
 DR PRINTS: PR00237; GPCR_HODOPS.
 DR PRINTS: PR00019; LEURICH_RPT.
 DR SMART: SM00370; LRR; 4.
 DR SMART: SM00013; LRRNT; 1.
 DR SMART: SM00369; LRR_Typ; 5.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
 DR PROSITE: PS00262; G_PROTEIN_RECEP_F1_2; 1.
 DR G-protein coupled receptor; Transmembrane; signal; Glycoprotein;
 KW Repeat; Leucine-rich repeat.
 FT SIGNAL 1 24
 FT CHAIN 25 951
 FT DOMAIN 25 544
 FT TRANSSEM 545 565
 FT DOMAIN 566 575
 FT TRANSSEM 576 596
 FT DOMAIN 597 619
 FT TRANSSEM 620 640
 FT DOMAIN 641 661
 FT TRANSSEM 662 682
 FT TRANSSEM 683 703
 FT TRANSSEM 704 724
 FT TRANSSEM 725 756
 FT DOMAIN 757 777
 FT TRANSSEM 778 783
 FT DOMAIN 784 804
 FT TRANSSEM 805 951
 FT DOMAIN 951 951
 FT REPEAT 55 79
 FT REPEAT 81 103
 FT REPEAT 104 127
 FT REPEAT 128 151
 FT REPEAT 153 175
 FT REPEAT 176 199
 FT REPEAT 200 223
 FT REPEAT 225 247
 FT REPEAT 248 270
 FT REPEAT 272 294
 FT REPEAT 318 341
 FT REPEAT 342 366
 FT REPEAT 368 387
 FT REPEAT 388 411
 FT REPEAT 413 435
 FT DISULFID 618 693
 FT CARBOHYD 68 68
 FT CARBOHYD 188 188
 FT CARBOHYD 199 199
 FT CARBOHYD 294 294
 FT CARBOHYD 314 314
 FT CARBOHYD 505 505
 SO SEQUENCE 951 AA: 104138 MW: EDD56AC072123461 CRC64;
 Query Match 29.2%; Score 967; DB 1; Length 951;
 Best Local Similarity 39.0%; Pred. No. 1.4e-63;
 Matches 212; Conservative 85; Mismatches 195; Indels 52; Gaps 5;

OY 235 WATVLLSVLCNGVLLVFPAGAPAPLPPVKKVVGAGANTLTGICGLIASVDALTFQ 294
 DB 546 WIFLVALLFVLVILVFA SCSSLPASKFGLISVSNLMIMYIGILTFIDAVSWGR 604
 OY 295 FSEYGARWETGLGRATGFLAVLGSSESVLLTLTAQCSVSVCVAYKSPISGVSRA 354
 DB 605 FAERGIMWETSGCKVAGSLAVFSSSEAVFLTLTAVERSVFAKDLKHGKSHLRFOV 664
 OY 355 GVLGLALAGLAALPLASVGEYASPLCLPYAPPEGPALGFTVALVMNSFCPLVA 414
 DB 665 AALLLGAAGAAGCPPLHGGQYASPLCLPF--PTGETPSLGFVTLVTLNLSLAFLLMA 722
 OY 415 GAVIKYCDLPFGDEAVWDCAMRHWAMLPADGLYCPLAFISFASMLDLFVTEPAV 474
 DB 723 IYTLNLCLEKEDLSENSQSSVIRKVAWLFTNCIFCPVAFSFAPLTAISPEIM 782
 OY 475 KSVLLVLPPLACINPLIYLPNPHRDDLRRLPRAG----- 512
 DB 783 KSVTLFFPLPCLNPLVLPVFNPKEDMKLKRVRKRGSVSVISQGGCEODFY 842
 OY 513 -DSGPLVAAAGLEKSSCDSTQALVAFSDVLLPASAAGPPLGTYGFPSTVLSQ 571
 DB 843 YDCGMVSHLQGNLTVDCCESFLTKPVSKHLI-----KSHSCPVLTAASQ 890
 OY 572 QPGA 575
 DB 891 RPEA 894
 RESULT 5
 LGR4_HUMAN STANDARD; PRT; 951 AA.
 ID LGR4_HUMAN
 AC Q9BXBL; Q9NYDL;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Leucine-rich repeat-containing G protein-coupled receptor 4 precursor
 DE (G protein-coupled receptor 48).
 GN GPR48 OR LGR4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=21294803; PubMed=11401528;
 RA Loh E.D., Broussard S.R., Kolakowski L.F. Jr.;
 RT "Molecular characterization of a novel glycoprotein hormone
 RT Biochem. Biophys. Res. Commun. 282:757-764(2001).
 RL G-protein-coupled receptor."
 CC -1- FUNCTION: Orphan receptor.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Expressed in multiple steroidogenic tissues:
 CC placenta, ovary, testis and adrenal. Expressed also in spinal
 CC cord, thyroid, stomach, trachea, heart, pancreas, kidney, prostate
 CC and spleen.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -1- SIMILARITY: CONTAINS 15 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: AF346711; AAK31153.1;
 CC EMBL: AF346709; AAK31153.1; JOINED.
 CC EMBL: AF346710; AAK31153.1; JOINED.
 CC EMBL: AF257182; AAF68989.1; -.

Genew; HGNC:13299; GPR48.
 DR MIM; 606666; .
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Mterm.
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR003591; LRR_typ.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 15.
 DR Pfam: PF01462; LRRNT; 1.
 DR PRINTS: PR00237; GPCRHOOPS.
 DR PRINTS: PR00019; LEURICHRPT.
 DR SMART; SM00370; LRR; 6.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_typ; 15.
 DR PROSITE: PS00237; G_PROTEIN_RECPT_F1; 1; FALSE_NEG.
 DR PROSITE: PS0262; G_PROTEIN_RECPT_F1_2; 1.
 DR G-protein coupled receptor; Signal; Transmembrane; Glycoprotein;
 KW Repeat; Leucine-rich repeat.
 FT SIGNAL 1 24
 FT CHAIN 25 951
 FT DOMAIN 25 544
 FT TRANSMEM 545 565
 FT DOMAIN 566 575
 FT TRANSMEM 576 596
 FT DOMAIN 597 620
 FT TRANSMEM 621 641
 FT DOMAIN 642 661
 FT TRANSMEM 662 682
 FT DOMAIN 683 703
 FT TRANSMEM 704 724
 FT DOMAIN 725 756
 FT TRANSMEM 757 777
 FT DOMAIN 778 783
 FT TRANSMEM 784 804
 FT DOMAIN 805 951
 FT REPEAT 55 79
 FT REPEAT 81 103
 FT REPEAT 104 127
 FT REPEAT 128 151
 FT REPEAT 153 175
 FT REPEAT 176 199
 FT REPEAT 201 223
 FT REPEAT 224 247
 FT REPEAT 248 270
 FT REPEAT 272 294
 FT REPEAT 318 341
 FT REPEAT 342 366
 FT REPEAT 368 387
 FT REPEAT 388 411
 FT REPEAT 413 435
 FT DISULFID 436 688
 FT CARBOHYD 68 693
 FT CARBOHYD 69 694
 FT CARBOHYD 199 294
 FT CARBOHYD 314 314
 FT CARBOHYD 505 505
 FT CONFLICT 292 292
 FT CONFLICT 433 433
 FT CONFLICT 668 668
 FT SEQUENCE 951 AA; 104460 MW; 5E0C2DFCF22C0A1BB CRC64;

Query Match Best Local Similarity 28.9%; Score 957; DB 1; Length 951;
 Matches 211; Conservative 85; Mismatches 191; Indels 60; Gaps 9;

DB 379 QIRKGTFOGLSLILDLSSRLHLHHSRAFRATPLTNLDVSNELTSPTEGLINQ 438
 OY 114 LKLGNLALSAFESKDSFPRKRLILEVAYVQCCPYGMCASFKAQGEADLHLD--- 170

DB 439 LKLVGNFKLAKALAKDFVNRSLSVYAYVQCAFAGWCDY----ANLNTEDNSLQDSV 494
 OY 171 -EESKRPGLGLARQANENHYDODDELQLEMEDSKPHPSVCCSPPTPQPKCEYLFESWG 229
 DB 495 AOEKGTADAAANVTSLNEHESQI-----IHCTSTGAFKPCCEYLGSSM 540
 OY 230 IRLAVNAIVLTVLCNGLVLTVPAGPADLPVKEVYVGAAGANTLGTISGLLASVDA 289
 DB 541 IRLTVWFIFLVALEFNLVLTFTFASCTSLPSSKLEIFGLISVNLPMGIYTGITFLDA 599
 OY 290 LTFGQSEYGARWETGLCRATGFLAVLSEASVLTTLTAAVCCSVSVYRAGVSPSL 349
 DB 600 VSMGRFAEFIMWETGSGCVAAGFLAVFSSESAFLMLATVRSLSAKDKMKGNSNHL 659
 OY 350 GSVRAVGLGLALAGLAAALPLASVBYGASPLCLPYAPPEGAPALGFTVALVMNSFC 409
 DB 660 KQRRVALAFLGATVAGFPLFRHGEYSASPLCLPF--PTGETPLSGFTVTLVLSLA 717
 OY 410 FLVAVAGATIKLYCDLPRGFEAVWDCAMVRHVAWLIFADGLLYCPVAFISFASMLGFPV 469
 DB 718 FILMAVITYIKLVONLEKEDLSENSQSMKHYAMWLTFTCIEFCVPAFSPFLTAISI 777
 OY 470 TPEAVKSVLLVLPPLACLPPLLYLLENPHFRD-----LRLRPRAQDGGPLAAYAAGEL 525
 DB 778 SPEIMKSVTLIFPLPACINPLVLYEFNPKFKEDMKLRRYTKKSGSVSISOGGCL 837
 OY 526 EK-----SSCDSTQALVAFSDV---LILEASEAGRPGLTGYFPSTV 566
 DB 838 EODFYDDGMYSHLOGNLTVCDCBSFLITKPKVCKHL-----KSHSCPALA 885
 OY 567 LISQCP 573
 DB 886 VASQCP 892

RESULT 6
 FSHR MOUSE STANDARD; PRT; 692 AA.
 ID FSHR MOUSE
 AC P5378; 090WY8; Q9D4C2;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Follicle stimulating hormone receptor precursor (FSH-R) (foliitropin receptor).
 DE FSHR.
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SV; TISSUE=Testis;
 RA Tena-Sempere M., Manna P.R., Huhtaniemi I.T.;
 RT Molecular cloning of the mouse follicle stimulating hormone receptor complementary deoxyribonucleic acid: functional expression of alternatively spliced variants and receptor inactivation by a C566T transition in exon 7 of the coding sequence.
 RT Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
 RL [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA MEDLINE=21083660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Aizawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Saito T., Okazaki Y., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H., Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita I., Pesole G., Quackenbush J., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barin G., Blake J., Boffelli D., Bojunga N., Carinci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

FT	REPEAT	218	240		BY SIMILARITY.		(POTENTIAL).	.	
FT	DISELID	441	516		N-LINKED (GICNAC.	.)	(POTENTIAL).	.
FT	CARBOHYD	191	191		N-LINKED (GICNAC.	.)	(POTENTIAL).	.
FT	CARBOHYD	199	199		N-LINKED (GICNAC.	.)	(POTENTIAL).	.
FT	CARBOHYD	293	293		N-LINKED (GICNAC.	.)	(POTENTIAL).	.
FT	CONFLICT	436	436		Q -> K IN REF. 2)				
FT	SEQUENCE	692 AA:	77769 MW:		4B57229180553A4		CRC64:		
SO									
	Query Match		13.3%;		Score 440;	DB 1:	Length 692;		
	Best Local Similarity		24.7%;		Pred. No. 6.7e-25;				
	Matches 167;		Conservative		93;	Mismatches 235;	Indels 182;	Gaps 25	

[illegible]

Query Match	13 08	Score	13 08
SO	SEQUENCE	694 AA; 78004 MW; EF077C5B8C8A54 CRC64;	
FT	CARBOHYD	293	
FT	CARBOHYD	268	
FT	CARBOHYD	199	
FT	CARBOHYD	191	
FT	DISULFID	441	
FT	REPEAT	218	
FT	REPEAT	193	
FT	REPEAT	168	
FT	REPEAT	119	
FT	REPEAT	93	
FT	REPEAT	68	
FT	DOMAIN	630	
FT	TRANSMEM	608	
FT	TRANSMEM	597	
FT	TRANSMEM	573	
FT	TRANSMEM	550	
FT	TRANSMEM	528	
FT	TRANSMEM	508	
FT	TRANSMEM	485	
FT	TRANSMEM	465	
FT	TRANSMEM	443	
FT	TRANSMEM	421	
FT	TRANSMEM	400	
FT	TRANSMEM	387	
FT	TRANSMEM	366	
FT	TRANSMEM	345	
FT	TRANSMEM	324	
FT	TRANSMEM	303	
FT	TRANSMEM	282	
FT	TRANSMEM	261	
FT	TRANSMEM	240	
FT	TRANSMEM	219	
FT	TRANSMEM	198	
FT	TRANSMEM	177	
FT	TRANSMEM	156	
FT	TRANSMEM	135	
FT	TRANSMEM	114	
FT	TRANSMEM	93	
FT	TRANSMEM	72	
FT	TRANSMEM	51	
FT	TRANSMEM	30	
FT	TRANSMEM	9	
FT	TRANSMEM	-12	
FT	TRANSMEM	-33	
FT	TRANSMEM	-54	
FT	TRANSMEM	-75	
FT	TRANSMEM	-96	
FT	TRANSMEM	-117	
FT	TRANSMEM	-138	
FT	TRANSMEM	-159	
FT	TRANSMEM	-180	
FT	TRANSMEM	-201	
FT	TRANSMEM	-222	
FT	TRANSMEM	-243	
FT	TRANSMEM	-264	
FT	TRANSMEM	-285	
FT	TRANSMEM	-306	
FT	TRANSMEM	-327	
FT	TRANSMEM	-348	
FT	TRANSMEM	-369	
FT	TRANSMEM	-390	
FT	TRANSMEM	-411	
FT	TRANSMEM	-432	
FT	TRANSMEM	-453	
FT	TRANSMEM	-474	
FT	TRANSMEM	-495	
FT	TRANSMEM	-516	
FT	TRANSMEM	-537	
FT	TRANSMEM	-558	
FT	TRANSMEM	-579	
FT	TRANSMEM	-600	
FT	TRANSMEM	-621	
FT	TRANSMEM	-642	
FT	TRANSMEM	-663	
FT	TRANSMEM	-684	
FT	TRANSMEM	-705	
FT	TRANSMEM	-726	
FT	TRANSMEM	-747	
FT	TRANSMEM	-768	
FT	TRANSMEM	-789	
FT	TRANSMEM	-810	
FT	TRANSMEM	-831	
FT	TRANSMEM	-852	
FT	TRANSMEM	-873	
FT	TRANSMEM	-894	
FT	TRANSMEM	-915	
FT	TRANSMEM	-936	
FT	TRANSMEM	-957	
FT	TRANSMEM	-978	
FT	TRANSMEM	-999	
FT	TRANSMEM	-1020	
FT	TRANSMEM	-1041	
FT	TRANSMEM	-1062	
FT	TRANSMEM	-1083	
FT	TRANSMEM	-1104	
FT	TRANSMEM	-1125	
FT	TRANSMEM	-1146	
FT	TRANSMEM	-1167	
FT	TRANSMEM	-1188	
FT	TRANSMEM	-1209	
FT	TRANSMEM	-1230	
FT	TRANSMEM	-1251	
FT	TRANSMEM	-1272	
FT	TRANSMEM	-1293	
FT	TRANSMEM	-1314	
FT	TRANSMEM	-1335	
FT	TRANSMEM	-1356	
FT	TRANSM		

Best Local Similarity 25.28; Pred. No. 4.8e-24;
Matches 149; Conservative 94; Mismatches 253; Indels 95; Gaps 14

QY	16	SGIGSGTHASVERSGSLTPAHPA-----SLALASNT-----TASGL-----	55
Db	92	SNLPKRIEIRKANNLLTYDHDAPQNLPLQVLLISNGIKHLPVHKTIQSLQVLLDI	151
QY	56	-----EXPFFSLS-----SLQALDSMAISIRPEAF-----STLHSL--	90
Db	152	QDNINITHYERNSEFMLESFESTLRIKSKNGIEIHNCACFNQTQDELNLSTNNNEELPN	211
QY	91	-----VKDLDTNQLTTLPLAGLGGMHLKLGKGNLALSQASXKSPFKRLTEVPY	141
Db	212	DYFQSGAGVILIDISQTRHSLPNYGLNLEKTLRARSTYNNLKLPSELEKVALMEANLTY	271
QY	142	AYQCCPYGMCASFRRKASG-----QHEADLHLDESSRRLGLLAR	183
Db	272	PSHCACAF---AMRRQTSSELQTCNKSLTRQEDVMTQARGSRVSLADDESSYKPGFDM	328
QY	184	QAEINHYDDLDDELQEMDSKPHSVQCCSPPPGPFKCEYLFEISWGTRLVWMAIVLLSVL	243
Db	329	YSEFEYD-----LCNEVVD-----VTSKPPDAFNCEDIMGDILRLVLPWISTIAIT	377
QY	244	CNGVILLTVFVGAPRLPEVKEVVGAIAGANTLTGISCGLASVDALTEQSEYCARWE	303
Db	378	GN-IIVILITTSOYKLTVPFLMCNLAFADLCIGITLLILASVDIHTISQYHNVAIDQ	436
QY	304	TGLCCRATGFLAVLGSEASVLLTLLAAVQCSVSCVRATGSPSLGSVRAGVLCGLAA	363
Db	437	TGAGCCDAAGFFVFASRLSYTLTATLERWHITTHAMOLECKVQDRIHAASVMLGWIFA	496
QY	364	GLAAALPLASVSGEYGASPLCPAPAPREGPALGTVALVMNNSCFELVAGAYKLYCD	423
Db	497	FAVALLPFGISITMYKVSICLPP--DISPSQLVMSLTVNLVLAFAVIGCYIHILY	554
QY	424	LPRGDE-LEAVMCCANRHYAHLIFADGLYCPAFVLFASPMGLFPYPEAVKSVLLVYL	482
Db	555	VNRPNVSSSTFKAKRWAILITFDICMAPISFFAISAKLPLTVTSKILLVLFY	614
QY	483	PLPACINPLLYLLENPHFRDRLRLPRPAGSGPLATNAAGELKSSCDST	533
Db	615	PINSCANPFLYALFTKFRDFILLSKFG-----CYEMQALVTRTST	660

RESULT 8
FSHR_BOVIN
ID FSHR_BOVIN STANDARD; PRT; 695 AA.
AC P35376;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Follltropin receptor).
GN FSHR.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
OC NCBI_Taxid=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Holstein; TISSUE=Ovary, and Testis;
RX MEDLINE=95127199; PubMed=7826612;
RA Houde A., Lambert A., Saumande J., Silversides D.W., Lussier J.G.;
RT "Structure of the bovine follicle-stimulating hormone receptor
complementary DNA and expression in bovine tissues.";
RL Mol. Reprod. Dev. 39:127-135(1994).
CC -I- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
ADENYLATE CYCLASE.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC FSH/LSH/TSB SUBFAMILY.

CC -!- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC between the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: L22319; AAC37324.1; -
 CC HSSP: P23945; 1XUN.
 CC InterPro: IPR000276; GPCR_Rhodopsn.
 CC InterPro: IPR001611; LRR.
 CC InterPro: IPR000372; LRR_Nterm.
 CC Pfam: PF00001; 7tm_1; 1.
 CC Pfam: PF00560; LRR_4.
 CC Pfam: PF01462; LRRNT; 1.
 CC SMART: SM00013; LRRNT; 1.
 CC PROSITE: PS00237; G-PROTEIN_RECEP_F1_1; 1.
 CC PROSITE: PS0262; G-PROTEIN_RECEP_F1_2; 1.
 CC G-protein coupled receptor; Transmembrane; glycoprotein; signal;
 CC Phosphorylation; Repeat; Leucine-rich repeat.
 KW SIGNAL; 17
 FT SIGNAL 1 695
 FT CHAIN 18 695
 FT DOMAIN 18 366
 FT TRANSSEM 367 387
 FT DOMAIN 388 398
 FT TRANSSEM 399 421
 FT DOMAIN 422 443
 FT TRANSSEM 444 465
 FT DOMAIN 466 485
 FT TRANSSEM 486 508
 FT DOMAIN 509 528
 FT TRANSSEM 529 550
 FT DOMAIN 551 573
 FT TRANSSEM 574 597
 FT DOMAIN 598 608
 FT TRANSSEM 609 630
 FT DOMAIN 631 695
 FT REPEAT 44 68
 FT REPEAT 69 93
 FT REPEAT 119 143
 FT REPEAT 170 192
 FT REPEAT 193 216
 FT REPEAT 218 240
 FT DISULFID 442 517
 FT CARBOHYD 191 191
 FT CARBOHYD 199 199
 FT CARBOHYD 293 293
 SQ SEQUENCE 695 AA; 78084 MW; 18F9DFEFC046380D CRC64;
 Query Match 12.8%; Score 424.5; DB 1; Length 695;
 Best Local Similarity 26.1%; Pred. No. 9,4e-24;
 Matches 135; Conservative 86; Mismatches 242; Indels 55; Gaps 10;

DB 391 OYKLVPEFELMGNLAFADLCIGIYLLLIASVDVHTKTEYHNHVAIDMOTGAGCDAAGFTV 450
 OY 317 LGSEASVLLTLTAAYOCVSVCYRAKSPSLGSVARGVGLAALGAAALPLASVGE 376
 DB 451 FASELSVTLTALTEHMTTHAMQLECKVQLRHAASIMLVGHTFPAVALFFIFGISS 510
 OY 377 YGASPLCIPYAPPEQAPALGFTVALVMNSFCFLVAGAVIKLYCDLPRGDF-EAYWDC 435
 DB 511 YKVSICLPM--DISPLSLQLYWMSLVNLVLAFAVIGCYTHILYVRNPNTSSSDT 568
 OY 436 AMRVHVALIFADGILLYCPVAFPLSFASMLGLEPPTPEAVKSVLVLPDPLNPLYTL 495
 DB 569 KIAKRAMAMLIETDFLCMAPISEFASISLKVPLTVSKILLVIFYFINSCANPLTAVI 628
 OY 496 FNPFRDRLRLPRAGDSGLVAAAGELKSSCDST 533
 DB 629 FTKNFRDRFILLSKFG-----CYEQVQATYSEHSST 661

RESULT 9
 FSHR_PIG STANDARD; PRT; 695 AA.

ID FSHR_PIG
 AC P49059; 077514;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin
 receptor).
 GN FSHR
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OC NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RX MEDLINE=96011644; Pubmed=7590277;
 RA Remy J.J., Labib-Mansais Y., Verle M., Bozon V., Couture L.,
 RA Payot E., Grebert D., Salesse R.;
 RT "The porcine follicle-stimulating hormone receptor: cDNA cloning, functional
 RT expression and chromosomal localization of the gene.";
 RT Gene 163:257-261(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RA Wang Y.F., Meyer K.B., Schmidt K., Wan S.J., Degen S.J.F.,
 RA la Barbera A.R.;
 RT "Porcine follicle-stimulating hormone receptor.";
 RT Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
 CC -!- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
 CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
 CC ADENYLATE CYCLASE.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC FSH/LSH/TSH SUBFAMILY.
 CC -!- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC between the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: L31966; AAA86933.1; -
 CC EMBL: AP025377; AAC24981.1; -
 CC HSSP: P23945; 1XUN.
 CC InterPro: IPR000276; GPCR_Rhodopsn.
 CC InterPro: IPR001611; LRR.
 CC InterPro: IPR000372; LRR_Nterm.
 CC Pfam: PF00001; 7tm_1; 1.

```

DR Pfam; PF00560; LRR; 2.
DR Pfam; PF01462; LRRNT; 1.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECIP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECIP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW phosphorylation; Repeat; leucine-rich repeat.
FT SIGNAL 1
FT CHAIN 17
FT DOMAIN 18 695
FT TRANSMEM 18 366
FT TRANSMEM 367 387
FT TRANSMEM 388 398
FT TRANSMEM 399 421
FT TRANSMEM 422 443
FT TRANSMEM 444 465
FT TRANSMEM 466 485
FT TRANSMEM 509 528
FT TRANSMEM 529 550
FT TRANSMEM 551 573
FT TRANSMEM 574 597
FT TRANSMEM 598 608
FT TRANSMEM 609 630
FT TRANSMEM 631 695
FT DOMAIN 44 68
FT REPEAT 69 93
FT REPEAT 119 143
FT REPEAT 170 192
FT REPEAT 193 216
FT REPEAT 218 240
FT DISULFID 442 517
FT CARBOHYD 191 191
FT CARBOHYD 199 199
FT CARBOHYD 293 293
FT CONFLICT 2 2
FT CONFLICT 13 13
FT CONFLICT 60 60
FT CONFLICT 166 166
FT CONFLICT 215 215
FT CONFLICT 247 247
FT CONFLICT 257 257
FT CONFLICT 334 334
FT CONFLICT 349 349
FT CONFLICT 352 352
FT CONFLICT 383 383
FT CONFLICT 407 407
FT CONFLICT 421 421
FT CONFLICT 427 427
FT CONFLICT 435 435
FT CONFLICT 483 483
FT CONFLICT 550 550
FT CONFLICT 586 586
FT CONFLICT 607 607
FT CONFLICT 691 691
SQ SEQUENCE 695 AA; 78172 MW; E9EBDB29C79C450 CRC64;

Query Match 12.84; Score 422; DB 1; Length 695;
Best Local Similarity 26.38; Pred. No. 1.4e-23;
Matches 137; Conservative 80; Mismatches 206; Indels 98; Gaps 12;

QY 55 LKXDTFSQLS-SIQALDLSWNAISHPPEAFS-----TLHSL----- 90
DB 160 VERNSEFGSLSESMILMLKNGREINHCAPNGQDELDLNSDNDNLEELPNDVFGASG 219
QY 91 -VVLDTLDNLTPLPLAGLGLMLHKLGNLALSOAFSKSPFKRLILEVYAVGCCPYG 149
DB 220 PYLLDISRRIRHSLPSYGLLENLKKLRKRAKSTYNNKLPLSLEKFTYLMASLTYSHCCAF- 278
QY 150 MCASFKAQSGWAEEDH-----LDDESSKRPLGLIARQ 184
DB 279 -----AMNRKISDPLHPLCKNSILRQEVYDVTGARGQVSLAEDEGESS-----LAKE 325
QY 185 AENHNDQDLDELQLEMEDESKPHPSVQCSPTPGPFKCEYLFEESWGLRLVAVLVLVLVC 244

```

```

DB 326 FDTWSEFYDLCNEVD-----YICSPDETFNCECEIMGHDLRLVIMFISILAIHG 379
QY 245 NCLVLLTFYAGGAPRLPVKRYVGAIGANTLTGISCGLASVDALTFQGFSEYGARMT 304
DB 380 N-IIVLVLLITSQYKLTVPKRFMLCMIAFDLCIGIYLLIASVDHTKQYHNVAIDMOT 438
QY 305 GLGCRATGEFLAVLGEASVLLTLTLA-----VQCSVSVSVAVYKSPSLGSV 352
DB 439 GAGCDAACFFTFYFASLSVYTLTLTLERMTTHAMOLCKV-----QLRHA 486
QY 353 RAGVIGTALAGLAALPLASVGEYASPLCLPYAPREGOPALGETVALVMNSFCFLV 412
DB 487 ASIMLVGWFIFETVALFPIFGISSYMKVSLCPM--DIDPSLSQLYVSVLVNVLAVV 544
QY 413 VAGAVIKYCOLPRGD--FEAVDCAWVHVMILFADGLYCVAFLSFAASMLGFVTP 471
DB 545 ICGCTHYTLTVRPNMSSSDTKRIAKRMAMLIETDFLCMAPISFALISASLKLPIIV 604
QY 472 EAVKSVLLVLPACLNPLLYLLENPFRRDLRLRPRAG 512
DB 605 SKSKILLVLEFPINSCANPLFYAFTKFRDVRDYLILSKRG 645

RESULT 10
LSHR_PIG
AC P16382; STANDARD; PRT; 696 AA.
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lutropin-choriogonadotropic hormone receptor precursor (LH/CG-R)
DE (LSH-R) (Luteinizing hormone receptor).
GN LHCGR.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
NCBI_Taxid=9823;
[1]
RX MEDLINE=89332517; PubMed=2502844;
RA Loosfelt H., Mistrall M., Alger M., Salese R., Thi M.T.V.H.-L.,
RA Jolivet A., Guichon-Mantel A., Sar S., Jallat B., Garnier J.,
RA Milgrom E.;
RT Cloning and sequencing of porcine LH-hcg receptor cDNA: variants
RT lacking transmembrane domain.*;
RL Science 245:525-528(1989).
CC -!- FUNCTION: RECEPTOR FOR LUTROPIN-CHORIOGONADOTROPIC HORMONE.
CC THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH
CC ACTIVATE ADENYLATE CYCLASE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- ALTERNATIVE PRODUCTS: 4 ISOFORMS; A (SHOWN HERE), B, C AND D; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC FSH/LSH/TSH SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DB EMBL; M29525; AAA31062.1; -
DR EMBL; M29526; AAA31063.1; -
DR EMBL; M29527; AAA31064.1; -
DR EMBL; M29528; AAA31065.1; -
DR PIR; A41344; A41344.
DR PIR; B41344; B41344.
DR PIR; C41344; C41344.
DR PIR; D41344; D41344.

```

DR HSSP: P22888; 1LUT.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Nterm.
 DR pfam: PF000001; 7tm.1; 1.
 DR pfam: PF000560; LRR_2.
 DR SMART: SM00013; LRRNT; 1.
 DR SMART: SM00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE: PS00262; G_PROTEIN_RECEP_F1_2; 1.
 DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_1; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing.
 FT SIGNAL 1 27
 FT CHAIN 28 696
 FT DOMAIN 28 358
 FT TRANSMEM 359 386
 FT DOMAIN 387 395
 FT TRANSMEM 396 418
 FT DOMAIN 419 439
 FT TRANSMEM 440 462
 FT DOMAIN 463 482
 FT TRANSMEM 483 505
 FT DOMAIN 506 525
 FT TRANSMEM 526 547
 FT DOMAIN 548 570
 FT TRANSMEM 571 594
 FT DOMAIN 595 605
 FT TRANSMEM 606 626
 FT DOMAIN 627 696
 FT REPEAT 122 147
 FT REPEAT 148 171
 FT REPEAT 172 196
 FT REPEAT 198 220
 FT REPEAT 221 244
 FT DISULFID 439 514
 FT CARBOHYD 99 99
 FT CARBOHYD 174 174
 FT CARBOHYD 195 195
 FT CARBOHYD 291 291
 FT CARBOHYD 299 299
 FT CARBOHYD 313 313
 FT VARSPLIC 317 329
 FT VARSPLIC 330 696
 FT VARSPLIC 317 331
 FT VARSPLIC 332 696
 FT VARSPLIC 317 628
 FT SEQUENCE 696 AA; 78092 MW; 593DEF1C25F982FE CRC64;
 Query Match 12.7%; Score 419; DB 1; Length 696;
 Best Local Similarity 25.28; Pred. No. 2.4e-23;
 Matches 134; Conservative 96; Mismatches 244; Indels 58; Gaps 12;
 QY 26 VERSGSLSPAPASIALAASNTASGKLEXTDFESQSS-----LOADLSMNA-IR 77
 DB 153 LEICDMLHTTYPAN-AFGGMNNESTLTKYNGFEIQSHAFNGTTLTSLKKNNAHLK 211
 QY 78 SHHPEAFSLHSLVKLDLTNDQTLTLAGLGIMHLKLGKMLATSOAFSKSPFKRLTL 137
 DB 212 KHNDAFRARGPSILDISSTRKLOALPSYGLSIOPLTITSYSKLKLPKRREKFTNLDA 271
 QY 138 EYFYAVQCCPY-----GMCASFKASQWAEEDLHLDDESSKRPGLGLAKQAEHY 189
 DB 272 TLTYPSHCCFNRLLPTKEQNFSSIRK-----NFSKQCESTARP-----NNETLY 317
 QY 190 -----DQDDELQLEWEDSKPHPSVQCSPTPGKPCCEYLFEEWG:RLAWAIVLLSVLC 244
 DB 318 SAIRFAESELDMWDYDGFCSF-KTLOCAPEPDAPNPEDIMGDFLRVLIMLILNIALMG 376
 QY 245 NGVLVLTVPFAGGAPAPLPVKEFVGAIGANTLTGISCGLLASVDALTEGOFSEYGARWET 304

DB 377 NYTVLFLVLTSHYKLTVP-RELMCNLSFADFCMGLYLLILASVDAQKRGQYXNAIDMQT 435
 QY 305 GLGCRATGFLAALGSEASVYLTLTAAVOCVSVSCVRAVGRKSPSGVAGVGLCALAG 364
 DB 436 GNGCSVAGFTVPASELSYTLTTLTLEWHHTITVAIQDQRLRHAIPIMLGWLFST 495
 QY 365 LAAALPLASVGEYASPLCLPYAPPEGOPALGFYVALVMNMSFCEIVAGAYIKLYCDL 424
 DB 496 LIAMPLVGVSSYMKVSIQLPM--DVETTLISQVYITLITLILNVVAFIICACIKITFAY 553
 QY 425 PRGDFEAV-WDCAMVHVAVMLTFADGLVCPVAFLSFASMLGLFPYTPRAVSVLLVLP 483
 DB 554 QNEPLMATNKDTRIAKMVAVLTDTTCAPAPISFPAISALAKVPLITVNSKVLVLFYF 613
 QY 484 LPACLPNPLVLFNPHFRDRLRLRPAGDSGPLAVAAAGELEKSSCDSTQA 535
 DB 614 VNSCANPFLAIFTKAFRRDFLL-----LSKSGCKRHOA 648
 RESULT 11
 FSHR_SHEEP STANDARD; PRT; 695 AA.
 ID FSHR_SHEEP STANDARD; PRT; 695 AA.
 AC P35379: Q28573; Q28574; Q9T819;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin receptor).
 GN FSHR.
 OS Ovis aries (sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Caprinae; Ovis.
 OC NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM FSH-R1).
 RC TISSUE=Testis;
 RX MEDLINE=93351750; PubMed=8394255;
 RA Yarney T.A., Sairam M.R., Khan H., Ravindranath N., Payne S., Seidah N.G.,
 RT "Molecular cloning and expression of the ovine testicular follicle stimulating hormone receptor".
 RL Mol. Cell. Endocrinol. 93:219-226(1993).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS FSH-R4 AND FSH-R3).
 RC STRAIN=Dorset-Leicester-Suffolk 1; TISSUE=Testis;
 RX MEDLINE=93176195; PubMed=8439338;
 RA Khan H., Yarney T.A., Sairam M.R.,
 RT "Cloning of alternately spliced mRNA transcripts coding for variants of ovine testicular follicle receptor lacking the G protein coupling domains".
 RL Biochem. Biophys. Res. Commun. 190:888-894(1993).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM FSH-R2), AND CHARACTERIZATION.
 RC STRAIN=Dorset-Leicester-Suffolk 1; TISSUE=Testis;
 RX MEDLINE=98031015; PubMed=10527886;
 RA Yarney T.A., Jiang L., Khan H., MacDonald E.A., Laird D.W., Sairam M.R.,
 RT "Molecular cloning, structure, and expression of a testicular follicle receptor with selective alteration in the carboxy terminus that affects signaling function".
 RL Mol. Reprod. Dev. 48:458-470(1997).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM FSH-R3), AND CHARACTERIZATION.
 RC STRAIN=Dorset-Leicester-Suffolk 1; TISSUE=Ovary;
 RX MEDLINE=20391225; PubMed=10527886;
 RA Babu P.S., Jiang L., Sairam A.M., Touyz R.M., Sairam M.R.,
 RT "Structural features and expression of an alternatively spliced growth factor type I receptor for follicle signaling in the developing ovary".
 RL Mol. Cell Biol. Res. Commun. 2:21-27(1999).
 RN [5]
 RP SEQUENCE OF 1-51 FROM N.A.

RX MEDLINE-98031017; PubMed-9364442;
 RA Saltram M.R., Subbarayan V.S.R.;
 RT "Characterization of the 5' flanking region and potential control
 RL elements of the ovine follicle stimulating hormone gene";
 CC MoJ. Reprod. Dev. 48:480-487(1997).
 CC -1- FUNCTION: Receptor for follicle stimulating hormone. The activity
 CC of isoform FSH-R1 is mediated by G proteins which activate
 CC adenylyl cyclase. Isoforms FSH-R2 and FSH-R3 also bind FSH, but
 CC this does not result in activation of adenylyl cyclase. Isoform
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Plasma membrane
 CC (isoforms FSH-R1 and FSH-R2); Cell surface (isoform FSH-R3).
 CC -1- ALTERNATIVE PRODUCTS: 4 isoforms; FSH-R1 (shown here), FSH-R2,
 CC FSH-R3 and FSH-R4, are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: Isoform FSH-R3 is expressed in ovary and
 CC testis, but not in kidney.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC FSH/LSH/FSH SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL, L07302; AAA31525.1; -
 DR EMBL, L12766; AAA31523.1; -
 DR EMBL, L12767; AAA31524.1; -
 DR EMBL, L36115; AAK70667.1; -
 DR EMBL, AJ31735; CA10495.1; -
 DR EMBL, AF090438; AAC61749.1; -
 DR PIR, JCI493; JCI493.
 DR HSSP; P23945; 1XUN.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Nterm.
 DR Pfam; PF00001; 7tm_1; 1.
 DR Pfam; PF01462; LRRNT; 1.
 DR SMART; SM00013; LRRNT; 1.
 DR PROSITE; PS00237; G-PROTEIN_RECPT; 1.
 DR PROSITE; PS00262; G-PROTEIN_RECPT; 1.
 KM G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KM Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing.
 FT CHAIN 1
 FT 18 695
 FT 18 366
 FT TRANSMEM 367 387
 FT DOMAIN 388 398
 FT TRANSMEM 399 421
 FT DOMAIN 422 443
 FT TRANSMEM 444 465
 FT DOMAIN 466 485
 FT TRANSMEM 486 508
 FT DOMAIN 509 528
 FT TRANSMEM 529 550
 FT DOMAIN 551 573
 FT TRANSMEM 574 597
 FT DOMAIN 598 608
 FT TRANSMEM 609 630
 FT DOMAIN 631 695
 FT REPEAT 44 68
 FT REPEAT 69 93
 FT REPEAT 119 143
 FT REPEAT 170 192
 FT REPEAT 193 216
 FT REPEAT 218 240
 FT DISULFID 442 517
 FT CARBOHYD 191 191
 FT CARBOHYD 199 199
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPPLIC 126 133 LISTNGIK -> FKRWRRI (IN ISOFORM FSH-R4).
 FT VARSPPLIC 135 695 MISSING (IN ISOFORM FSH-R4).
 FT VARSPPLIC 224 259 DISRRIRSLPSYGLLENKRLAKSTYHLKPSLE -> SPLLCMAHLOSFFVVGOREHISEGLSKOHNP (IN ISOFORM FSH-R3).
 FT VARSPPLIC 260 695 MISSING (IN ISOFORM FSH-R3).
 FT VARSPPLIC 643 670 KFGCEYVOAQYRSESTFANFHPRNG -> LHCCNGLI CNHESSELVARGNIFLND (IN ISOFORM FSH-R2).
 FT VARSPPLIC 671 695 MISSING (IN ISOFORM FSH-R2).
 SQ SEQUENCE 695 AA; 78237 MW; FEF75D89D86C0DBA CXC64;
 Query Match 12.7%; Score 418.5; DB 1; Length 695;
 Best Local Similarity 26.3%; Pred. No. 2.6e-23;
 Matches 139; Conservative 80; Mismatches 225; Indels 85; Gaps 13;
 QY 55 LEXDIFSQLS-SLOALDISMAIRSIHPEAF-----STLHSU----- 90
 Db 160 VERNSEFMGLSESMIWLKNGIOEIHNCANQDDELNLSDNSLLELPNDVFGAGS 219
 QY 91 -VKLDITDNOITPLIAGLGIMHLKGNLALSOFKSPKRLILEVPAVQCCPYG 149
 Db 220 PYLIDISTRIRSLPSYGLLENKRLAKSTYHLKPSLEKFTLVPAASLTYPSCCAF- 278
 QY 150 MCASFRRASGQWEADLH-----LDDESSKRPGLLARQAEHNYDDDELQ 198
 Db 279 -----ANMRQTSDDHPICNKSILRQEVDMQARGQISLAEDDESYAKGFMKYS 331
 QY 199 EMEDS--KPHSVQCSPTPPEKPECELFESWGRILAWAIVLVSVLCNGVLTTFVAG 256
 Db 332 EFDYDLCESEVVDYTCSEPEPAFPCEDIMGYDILRLVLEFISLITATGNLVLV- 390
 QY 257 PAPLPYKKEVVAIAGANTLTGICGLASVDALTGQSESGEAGRWETGCRAGFLAV 316
 Db 391 QYRLVPRFLMCNMFADLCIGYLLIASVDHTSQYHNVAIDMVGAGDAAGFFTV 450
 QY 317 LGSEASVLLTLTAAVO-----CSVASC-----VRAGKSPSGSVRAGVGLAGLAA 367
 Db 451 FASELSVYTLTATLEFRMTTHHMOLECKVHRIHAASIMLVGNV-----FAFAVA 501
 QY 368 ALPIASVGEYASPLCEPAPPEGQPAALGFVALVMMNSPCFLVAVAGATIKYDLP 427
 Db 502 LPEIFDISMYKMSVLCIPW--DIDSPLSQLYMSLVLNVLAFFVIGCCYTHIYVNP 559
 QY 428 DF-EAVWDCAWVRHVAWLLFADGLIXCYAFLSFSMLGLEPVTPAVKSVLVLP 486
 Db 560 NITSSSDPKIKKRAMMLFTDFLCMAPISEFAISASIKVPLTVSKSLDLVLFYINS 619
 QY 487 CLNPLLYLLEHFRDRLRLPRAGDSGPLAVAAAGLEKSSGSDSTQA 535
 Db 620 CANPLLYAIFTRNFRDFEL-----LSKFGCEYVOA 651
 RESULT 12
 FSHR_RAT
 ID FSHR_RAT STANDARD; PRT; 692 AA.
 AC P20395;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Follicle stimulating hormone receptor precursor (FSH-R) (Follicle stimulating hormone receptor).
 GN FSHR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 ON NCBI_TaxID=10116;
 RX [1] SEQUENCE FROM N.A. TISSUE=seroli cells; MEDLINE-91125358; PubMed-2126341; Sprengel R., Braun T., Nikolics K., Segaloff D.L., Seeburg P.H.;

*The testicular receptor for follicle stimulating hormone: structure
RT and functional expression of cloned cDNA.*
Mol. Endocrinol. 4:525-530(1990).

CC EMBL; L02842; AAA1175.1; -

CC PIR; A34548; A34548.

CC PIR; A41729; A41729.

CC HSSP; P23945; 1XUN.

CC Interpro: IPR000276; GPCR_Rhodopsin.

CC Interpro: IPR001611; LRR.

CC Interpro: IPR000372; LRR_Nterm.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

CC PIR; A41729; A41729.

Matches 167; Conservative 86; Mismatches 251; Indels 164; Gaps 23;

16 SGIPGSHASVERSGG-----SLPAPASIAAALASNTASGKLEXPFSQSL 66

92 SNLPLHEIRLERKANNLLINPAFONLP-----SLRYLLISNT-----GIKHPAYHKIOSL 144

67 Q-----ALDSMNAIRSIHPEAF-----S 85

145 QVLLDIODININIHVARNSEFGISFESVILNLSKNGIEIHNCACFNGTDELNLSDN 204

86 TLHSL-----VLDLDTNQLTTLPLAGLGMLKTKGNALSOAFSKDSFPKL 134

205 NLELPNDVPOGASCPVLLDISRTKVSHPNHLGLENLKTIRASTRYLKLPLDKFVTL 264

135 RILEVYAYOCPCYGMCAFFKASGOWEADLH-----LD 169

265 MEASLTYPHSCAF-----ANLKRQISELHPICNKSILRODIDMTIGORVSLD 316

170 DESSKRPGLGLARQENHYDDDELQLEMDSDSKPHPSVOCSPTPGPFKCEYLFSWG 229

317 DEFS-----YKGSDDMYNEDYDLN-----EYVD-----VTCSPKPDAFNCEIDMGYNT 363

230 IRLAVNAVLIVLNCGLVLLTFVAGPAPLPVKFVGATAGNTLIGISGLASVDA 289

364 LRVLLWFSLIATIGNTVLV-VLTSSQYKLTVPRLMCLNLFADLCIGIYLLIASVDI 422

290 LFFGPFSEYGARWETGLGCRATGFLAYLGEASVLLTLAAVQ-----CSVSYCVRAY 343

423 HRKSOYHNYAIDWGTGACCDAGFFYFASBELSVYTLTALTEHMTTHAQLDEC 478

344 GKSPSLGVRAGVLCGLALAGLAALPLASVGEYGASPLCPYAPREGPALGFTVALV 403

479 -KVQLRHASVWVLG-WTFARFAALFPIFGISYMKVSLCPM--DIDPSLSQLYMAL 534

404 MNSFCEFLVAGAYIKIVCDLPGRDF-EAVWDCAVHRVAMLIAPADLLCYPAFLSPAS 462

535 VLVNVAEVAIVGCTHYITITVNPPLTVSSSDTKIAKMAVLIITDLCAPISFAISA 594

463 MGLFPTPEAVKSVLLVLPACINPLLYLFNPHRDDLRLRLRPAGDSGLAVAA 522

595 SLKPLPLTVSKAKILVLEFPINSCANPLVLAIFTKNRRODFIL-----639

523 GELKSCSDSTQALVARSVDLLLEASGAPPGLETFYFVTLISCOGAPRLSEGS 582

640 -LSKFGCYEMQAOIYNT-----ETISSA-----TNF-HAKSHCS--SAPVTSY 681

583 CYEPENH 590

682 VLVPL-NH 688

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

23;

66

144

85

204

134

264

169

316

229

363

289

422

343

478

403

534

462

594

522

639

582

681

688

688

688

688

688

688

688

688

688

688

688

688

688

688

688

23;

66

144

85

204

134

264

169

316

229

363

289

422

343

478

403

534

462

594

522

639

582

681

688

688

688

688

688

688

688

688

688

688

688

688

688

688

688

RP SEQUENCE FROM N.A.
 RC TISSUE-Testis;
 RA MEDLINE=93246012; PubMed=1301382;
 RA Kelton C.A., Cheng S.V., Nugent N.P., Schmelckhardt R.L.,
 RA Rosenthal J.L., Overton S.A., Wands G.D., Kuzesja J.B., Luchette C.A.,
 RA Chappel S.C.;
 RT "The cloning of the human follicle stimulating hormone receptor and
 RT its expression in COS-7, CHO, and Y-1 cells.";
 RL Mol. Cell. Endocrinol. 89:141-151(1992).
 RP SEQUENCE FROM N.A.
 RA Tilly L.T., Alhara T., Nishimori K., Jai X.-C., Billig H.,
 RA Kowalski K.I., Perlas E.A., Hsueh A.J.;
 RL Submitted (xxx-1992) to the EMBL/Genbank/DBJ databases.
 RP SEQUENCE OF 1-342 FROM N.A.
 RC TISSUE-Testis;
 RA MEDLINE=93075197; PubMed=1359889;
 RA Gromoll J., Gudermand T., Nieschlag E.;
 RT "Molecular cloning of a truncated isoform of the human follicle
 RT stimulating hormone receptor.";
 RL Biochem. Biophys. Res. Commun. 188:1077-1083(1992).
 RP SEQUENCE OF 1-51 FROM N.A.
 RA MEDLINE=95011044; PubMed=7926278;
 RA Gromoll J., Dankbar B., Gudermand T.;
 RT "Characterization of the 5' flanking region of the human follicle-
 RT stimulating hormone receptor gene.";
 RL Mol. Cell. Endocrinol. 102:93-102(1994).
 RP 3D STRUCTURE MODELING OF 49-228.
 RA MEDLINE=96363672; PubMed=8747461;
 RA Jiang X., Dreano M., Buckler D.R., Cheng S., Ythier A., Wu H.,
 RA Hendrickson W.A., El Tayar N.;
 RT "Structural predictions for the ligand-binding region of glycoprotein
 RT hormone receptors and for the nature of hormone-receptor interactions";
 RL Structure 3:1341-1353(1995).
 CC -1- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
 CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
 CC ADENYLATE CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: SEROTOLI CELLS AND OVARIAN GRANULOSA CELLS.
 CC FSH/LSH/TSH SUBFAMILY.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -1- SIMILARITY: CONTAINS 5 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: M65085; AAAS2477.1;
 DR EMBL: S59800; AAB26480.1;
 DR EMBL: M95489; AAAS2478.1;
 DR EMBL: X68044; CAA48179.1;
 DR EMBL: S73199; AAB32071.1;
 DR PIR: JN0122; JN0122;
 DR PDB: 1XUN; 15-MAY-97.
 DR Genew; HGNC:3969; FSHR.
 DR MIM: 136435;
 DR Interpro: IPR000276; GPCR_Rhodopsn.
 DR Interpro: IPR001611; LRR.
 DR Interpro: IPR000372; LRR_Nterm.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 3.
 DR Pfam: PF01462; LRRNT; 1.
 DR SMART: SM00013; LRRNT; 1.
 DR PROSITE: PS00237; G_PROTEIN_RECPT_FL_1; 1.

DR PROSITE: PS0262; G-PROTEIN_RECPT_FL_2; 1.
 KW G-Protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing;
 KW 3D-structure.
 FT SIGNAL 1 17
 FT CHAIN 18 695
 FT DOMAIN 18 366
 FT TRANSMEM 367 387
 FT TRANSMEM 388 398
 FT TRANSMEM 399 421
 FT TRANSMEM 422 443
 FT TRANSMEM 444 465
 FT TRANSMEM 466 485
 FT TRANSMEM 486 508
 FT TRANSMEM 509 528
 FT TRANSMEM 529 550
 FT TRANSMEM 551 573
 FT TRANSMEM 574 597
 FT TRANSMEM 598 608
 FT TRANSMEM 609 630
 FT TRANSMEM 631 695
 FT REPEAT 69 93
 FT REPEAT 119 143
 FT REPEAT 170 192
 FT REPEAT 193 216
 FT REPEAT 218 240
 FT DISULFID 442 517
 FT CARBOHYD 191 191
 FT CARBOHYD 199 199
 FT CARBOHYD 293 293
 FT CARBOHYD 318 318
 FT VARSPLIC 224 285
 FT VARSPLIC 342 695
 FT CONFLICT 13 13
 FT CONFLICT 112 112
 FT CONFLICT 197 198
 FT CONFLICT 295 295
 FT CONFLICT 307 307
 FT CONFLICT 680 680
 SO SEQUENCE 695 AA; 78294 MW; 723B8E1F76DCD5 CRC64;
 Query Match 12.5%; Score 414.5; DB 1; Length 695;
 Best Local Similarity 25.0%; Pred. No. 5.1e-23;
 Matches 131; Conservative 94; Mismatches 232; Indels 67; Gaps 13;
 QY 55 LEXDTPSOIS-SLOALDSWMAIRSIHPEA-----STLHSL----- 90
 DB 160 IERNSPVGLSFESVILIMLNKNGIOEIHNCANFGQGLDELNSDNNNLEBPDVFGASG 219
 QY 91 -VKLDLTNQLTTLPLAGLIGLMLKTLKGNLALSOAHSKSPFKRLILEVPAVYCCPYG 149
 DB 220 PYLIDISRTIRHSISYGLLENMKTLRASSTYNLKLPTLRLVALMASLTPSHCCAF- 278
 QY 150 MCASFYKASGQWEAEDH-----LDDESSRRPLGLARQAEHNYDDDELQ 198
 DB 279 -----AMNRQISLHPLPCNKSITLROEVDYMTQTRGSSLAEDNESSYRGDMYTT 331
 QY 199 EMDSKRHP--SVQSPFPKPFKCEYLFESGIRLAWAIVLSYLCGLVLTVFAGG 256
 DB 332 EFDYDLONEVVDYCSKPKPAFNPCEIDINGINILKILWFISILAITGN-IIVILTTS 390
 QY 257 PAPLPYKPVVGAAGANLTLTSGGLASVALTFGQFSEYGARWETGTCRATGLAY 316
 DB 391 QYKLTVPRFLMCLNAPADICITIVILLASVDIHRKSOYHNVAIDMOTGACGDAAGFTV 450
 QY 317 IGSEASVLLTLTAAYO-----CSVSVCAVAYGKSPSLGVRAGVAGCGLAAGLAALP 370
 DB 451 FASELSVYTLTILTERMHTITHAMQDLC-----KYDLRAHASVYVWGVI-FAFAALFP 504
 QY 371 LASVEGYASPLCLPYAPPEGQPAALGFTVALVMAHNSFCGLVAVAGAIKYLCDLPFGDF- 429
 DB 505 IFGISSYKVSICLPM--DIDSPLSQLYVMSLVLNVLAVVICGYIHLYLVRRNPNI 562

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 NCBI_TaxID=9615;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thyroid;
 RX MEDLINE=90098886; PubMed=2602159;
 RA Parentier M., Libert F., Maenhaut C., Lefort A., Gerard C.,
 RT Perret J., van Sande J., Dumont J.E., Vassart G.,
 RL "Nucleotide sequence of the dog thyrotropin receptor cDNA."
 Nucleic Acids Res. 17:10493-10493(1989).
 [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thyroid;
 RX MEDLINE=90084524; PubMed=2556796;
 RA Parentier M., Libert F., Maenhaut C., Lefort A., Gerard C.,
 RT Perret J., van Sande J., Dumont J.E., Vassart G.,
 RL "Molecular cloning of the thyrotropin receptor."
 Science 246:1620-1622(1989).
 CC -1- FUNCTION: RECEPTOR FOR THYROTROPIN. PLAYS A CENTRAL ROLE IN
 CONTROLLING THYROID CELL METABOLISM. THE ACTIVITY OF THIS
 RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE
 CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -1- FSH/LSH/TSH SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 4 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X17146; CAA35026.1; -
 DR EMBL: X17147; CAA35027.1; -
 DR EMBL: M29957; AAA30901.1; -
 DR EMBL: M90047; AAA30902.1; -
 DR PIR: S06933; S06933.
 DR PIR: A40077; A40077.
 DR HSSP: P16473; 1XUM.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR Pfam: PF00560; LRR_1; 1.
 DR PROSITE: PS00237; G-PROTEIN_RECEPTOR_FL_1; 1.
 DR PROSITE: PS50262; G-PROTEIN_RECEPTOR_FL_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing.
 FT SIGNAL 1 20
 FT CHAIN 21 764
 FT DOMAIN 21 413
 FT TRANSMEM 414 441
 FT DOMAIN 442 450
 FT TRANSMEM 451 473
 FT DOMAIN 474 494
 FT TRANSMEM 495 517
 FT DOMAIN 518 537
 FT TRANSMEM 538 560
 FT DOMAIN 561 580
 FT TRANSMEM 581 602
 FT DOMAIN 603 625
 FT TRANSMEM 626 649
 FT DOMAIN 650 660
 FT TRANSMEM 661 682
 FT DOMAIN 683 764
 FT REPEAT 51 74
 FT REPEAT 150 174
 FT REPEAT 176 199
 FT REPEAT 201 223
 FT DISULFID 494 569
 BY SIMILARITY.

FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 81 105 MISSING (IN SHORT ISOFORM).
 SO SEQUENCE 764 AA; 86483 MW; 49F03B3DBC85512 CRC64;
 Query Match 12.5%; Score 414.5; DB 1; Length 764;
 Best Local Similarity 23.7%; Pred. NO. 5.7e-23;
 Matches 141; Conservative 102; Mismatches 258; Indels 95; Gaps 12;
 QY 33 SLPAHPASIALAASNTASGKLEDTFSQ-----SIQALDLSMNA-IRSHHPAF 84
 Db 166 SIPAN-----AFQGLCNEELTLKLVNNGTSTIQGAHFNCKLDAYLNKKYLSAIDKDAF 221
 QY 85 STLMS-LVKLDLTDNQLTTPLAGIGGLMLKGNLALSOAFSKDSFPKRIIEVPAY 143
 Db 222 GGVSSEPTLIDVSTYVALPSKGLHKLKELIARNTWTLKLPISLFLHLTRADLSYPS 281
 QY 144 QCCPY-----GMCASP-----FRASG 159
 Db 282 HCCAFKKQKKIRGLIESLMCNESIRSLRQKSVNTLNGPPDQYEYELGDSHAGYKDNS 341
 QY 160 QW-----BAEDLHLDDSSSKRPLGLARQAEHNDODLDELQLEMEDSKP 205
 Db 342 QFQDTDSNHHYVFEFEQDEEILQFQELKNPQETLQARDSHYDLY-----CGG 392
 QY 206 HPSVQSPPTGPKPKPEYLFESMGIRLAWAIVILSVLCGLVILYFAGGAPAPLPVKF 265
 Db 393 NEDMVCIRPKSDERFPCDINGYFELRLVWVFSLALGLVNFVILYLTSHYKLYP-RF 451
 QY 266 VQALAGANTLTISGGLASVDALTFQPSSEYGARWETGLGCRANGFLAVLSEASVYL 325
 Db 452 LMCNLAFDFPCMGMYLLILASVDLYTHSEYYNHAIDMGQTGPGCTNGFETVFEASELSVYT 511
 QY 326 LFLAAVQCSVSCVAVAYGSPSLGSRVAGVLCIALAGLAAALPLASVGEYASPLCLP 385
 Db 512 LTVITLERWATIFAMRLDKIRLRHAYVAIMVGWCCFLALLPLVGISYAKVSICLP 571
 QY 386 YAPESOPALGFTVALVMNSFCFLVAVAGAYIKLYCDLPFGDFE-AVMDCAVVRHVAWL 444
 Db 572 M--DTETPLALVILVLLINIVAFITVCSYKVIITVRNPQYNGDDTKIAKMAVL 629
 QY 445 IFADGLGCPVAFELSPASMLGLFPTVPEAVKSVLVLPPLPCLNPLLYLLENPHRDDL 504
 Db 630 ITDECMAPISFYALSALMKNPLITVNSKILLVLEFPLNSCANPFLYAITKFAQRDV 689
 QY 505 RLRLPRAGDSGLAYAAAG-----ELEKSSCDSTQALVAFSDVDLLEAS 549
 Db 690 FILLSKFGICKRQAGAYRGQVRSPKNSAGIQIQKTYTRDMKRSLLPMQDEYELLENS 745

Search completed: November 8, 2002, 19:34:53
 Job time : 14.2476 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2002, 19:32:55 : Search time 16.1833 Seconds
(without alignments)
3760.229 Million cell updates/sec

Title: US-09-851-595-5
Perfect score: 3307
Sequence: 1 NTHHESMYACRYRSGIPC.....GGLSGGGGPGPSGLAFASHV 633

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1275	38.6	907	2	JE0176 orphan G protein-c
2	1205	36.4	907	2	JE0193 G protein-coupled
3	429	13.0	694	2	JC4301 follicitropin recept
4	428.5	13.0	694	2	JC2237 follicitropin recept
5	424.5	12.8	695	2	A45896 follicle stimulat
6	419	12.7	696	2	A41344 lutropin-choriogon
7	418.5	12.7	695	2	JC4493 follicitropin recept
8	417	12.6	692	2	A34548 follicitropin recept
9	414.5	12.5	695	1	ORH007 follicitropin recept
10	414.5	12.5	695	1	JN0898 follicitropin recept
11	414.5	12.5	764	2	A40077 thyroid stimulat
12	414.5	12.5	764	2	JC5643 follicitropin recept
13	412.5	12.5	700	2	A42395 lutropin receptor
14	408	12.3	764	2	A48882 lutropin receptor
15	407.5	12.3	700	2	I77463 luteinizing hormon
16	407.5	12.3	700	2	A49744 lutropin-choriogon
17	403.5	12.2	764	1	ORH007 thyroid stimulat
18	400	12.1	814	2	JC7389 thyroid stimulat
19	397.5	12.0	793	2	JC7390 follicitropin recept
20	384.5	11.6	696	2	JC7361 follicitropin recept
21	381.5	11.5	764	2	A35956 thyroid stimulat
22	381	11.5	699	1	ORH007 lutropin-choriogon
23	366	11.1	889	2	J20123 hypothetrical prote
24	308.5	9.3	925	2	JC2033 G protein-coupled
25	299	9.0	1115	2	S40241 lysine carboxypept
26	149.5	4.5	536	2	A34901 iodopsin homolog -
27	137.5	4.2	355	1	A46191 hypothetrical prote
28	136.5	4.1	653	2	T25194 hypothetrical prote
29	131.5	4.0	1784	2	C96615

30	131	4.0	350	2	I3848	Mel-1a melatonin r
31	130	3.9	440	2	JC5520	serotonin receptor
32	130	3.9	440	2	A45121	alpha-1b adrenergi
33	130	3.9	864	2	T08575	protein kinase hom
34	129.5	3.9	367	2	JE0349	interferon-inducib
35	129.5	3.9	658	2	B04664	probable receptor-
36	129	3.9	605	2	A41915	insulin-like growt
37	128	3.9	366	2	I46469	hypothetrical prote
38	127.5	3.9	382	2	E84527	galanin receptor 2
39	127.5	3.9	387	2	JC5949	silt protein 2 pre
40	127.5	3.9	1469	2	B36665	silt protein 1 pre
41	127.5	3.9	1480	2	A36665	opsin, green-sensi
42	126.5	3.8	349	2	A45229	opsin, green-sensi
43	126.5	3.8	349	2	B45229	adrenomedullin rec
44	125.5	3.8	404	2	JC5784	insulin-like growt
45	125	3.8	603	2	JC6128	

ALIGNMENTS

RESULT 1

JE0176
orphan G protein-coupled receptor precursor - human
C:Species: Homo sapiens (man)
C:Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 21-Jul-2000
C:Accession: JE0176
R:McDonald, T.; Wang, R.; Bailey, W.; Xie, G.; Chen, F.; Caskey, C.T.; Liu, Q.
Biochem. Biophys. Res. Commun. 247, 266-270, 1998
A:Title: Identification and cloning of an orphan G protein-coupled receptor of the gl
A:Reference number: JE0176; MUID:98308104; PMID:9642114
A:Accession: JE0176
A:Molecule type: mRNA
A:Residues: 1-907 <MCD>
A:Cross-references: GB:AF062006; NID:q3366801; PIDN:AAC28019.1; PID:q3366802
A:Comment: This protein is a receptor for a novel class of glycoprotein ligands.
C:Genetics:
A:Gene: HG38
A:Map position: 12q22-23
F:1-21/Domain: signal sequence #status predicted <SIG>
F:562-583/Domain: transmembrane #status predicted <TM1>
F:594-616/Domain: transmembrane #status predicted <TM2>
F:639-660/Domain: transmembrane #status predicted <TM3>
F:681-701/Domain: transmembrane #status predicted <TM4>
F:725-744/Domain: transmembrane #status predicted <TM5>
F:768-791/Domain: transmembrane #status predicted <TM6>
F:803-824/Domain: transmembrane #status predicted <TM7>

Query Match 38.6% Score 1275; DB 2; Length 907;
Best Local Similarity 46.9% Pred. No. 1.2e-91;
Matches 250; Conservative 96; Mismatches 157; Indels 30; Gaps 8;

QY	54	KLEKDFPQSLSSALDLSNNATRSIHPEAFSTLSLVKLDLTTPPLAGLGGLH	113
DB	388	EIKVDFPQQLSLRSNLNANWKTAIIHPNFSPLPKIDLSNLSPPITGLHGLT	447
QY	114	LKIKGNLALSQAFSKDSFPLRLILEVPIYAOCCPYGMCASFRAASQW-----EAEEDH	167
DB	448	LKLTGNHAIQSLISSENPFLKVIEMPYAYOCCAFCEYAKRISNQWNKGDSSMDLH	507
QY	168	LDDEESKRLGLLQAQAEHNYDDQLEDEQLDEM-EDSKRPSQCSPTPGPKPCETLFE	226
DB	508	-----KKDGMFQAQDE---RDLEDFLDEEDIKALHSQCSFSPKPKCEILD	556
QY	227	SWGIRLAWAIVLISVLCNGLVLLTFVAGGAPDLPPYKVVGAIGANTLTGISGLLAS	286
DB	557	GWLIRIGWIRIVALLCNALVTSVFR-SPIYISPIKILIGYIAAVNMILTGVSSAVLAG	615
QY	287	VDALTFQSESEYARWETGICRATGFLAVGSEASVLLTTAAVQCSVSVCAVRAKGS	346
DB	616	VDATFGSEFARHGAWENGVCYVIGLSIFASESSVFLLTAALEGFVSKYSAKFEIK	675
QY	347	PSLGSVAGVGLGICALAGLAALPLASVGEYASPLCLPYAPPEGGPAALGFTVAIVMNM	406

Db 676 APFSSLKIIILLCALALTMMAVPLLGSGKYGASPLCLPL--PCEPSPMGWMLIILN 733
OY 407 SECFVAVGATIKLYCDLPRGDFEAVWDCAMVRHVAAILFADGLLYCYVALISFASMLGL 466
Db 734 SLCTLMATATATKLYCNDKGDLENIMDCSMVKHIALTLFNCILNCEVAFISFSSILNL 793
OY 467 FPVPEAVKSVLLVPLPACLNPLLYLFNPHFRDRLRPRRA-----GDSGFLAYAA 521
Db 794 TEISPEVAKFLVPLPACLNPLLYLFNPHFRDRLRPRRA-----GDSGFLAYAA 521
OY 522 AGELEKSSCDSTQALVAFSDVDLLEASGRP---PGLTYGFPVTLISC 570
Db 854 SDDVEKOSCDSTQALVFTSSITVDLPSSVSPSPAPVPSCHLSVAFYPC 906

RESULT 2

G protein-coupled receptor FEX - mouse
C/Species: Mus musculus (house mouse)
C/Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000
C/Accession: JG0193
R/Hermey, G.; Melhner, A.; Schaller, H.C.; Hermans-Borgmeyer, I.
Biochem. Biophys. Res. Commun. 254, 273-279, 1999
A:Title: Identification of a novel seven transmembrane receptor with homology to glycoprotein A:Reference number: JG0193; PMID:9920770
A:Accession: JG0193
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-907 <HBR>

Query Match Best Local Similarity 36.48; Score 1205; DB 2; Length 907;
Matches 242; Conservative 102; Mismatches 165; Indels 18; Gaps 8;

OY 54 KLEDFPSQSSLOALDLSNNAIRSIHPEAFSTHSYKIDLDNOLTTPLAGLGLMH 113
Db 388 EIGSTFQQLFNLRSILNANKIAIHPNAPSTLPSLIKIDLSNLSPPVGLHGLTH 447
OY 114 LKIKGNLALSOAFSKDSFPRKILEVYAYQCCPYGMCASFRRAGOWEADLHLDDEES 173
Db 448 LKLTGNRALSLPSANFPELKIEHMSAVQCAFSGCENYKISQNMNDDGNSVD-DL 506
OY 174 SKRPLGLANQAEHNDODLDELQLEK-EDSKRHPVSQCFPTGPRPCYLFESGIRL 232
Db 507 HKKDALFOYQDE---RDLEDFLLDFEEDLNAHVSQCSPPGPKPCHELGSMILRI 562
OY 233 AWAIVLLSVLNGVLLTVFAGGAPLPVVKFVGAIAAGANTLTGICGLASVADLT 292
Db 563 GWTATVAVLTSCNALVALVFER--TPLYISIKLIGIVAVDILMGVSSAVLAADAFTR 621
OY 293 GQSEVGAHWENGIGCRATGFLAVLGSEASVLLTTAAVQCSVSVSCVRAVYKSPSLGSV 352
Db 622 GRHOAGWAMWEDIGQIVGFLSIFASESSIFLTLTAALEGRFSVSCSKFEKAPLEFSL 681
OY 353 RAGVLEGLALAGLAAALPLASVEYGAASPLCLPYAPPEGAPALGPTVALVMMNSCEIV 412
Db 682 RAIVLLCVLALATLAIITPLIGSGKYASPLCLPL--PREPSTTGIMVALVNLNLCFTL 739
OY 413 VAGATIKLYCDLPRGDFEAVWDCAMVRHVAAILFADGLLYCYVALISFASMLGLFPVTP 472
Db 740 KMTAAVTKLYCSLEKGELEMLMCDMSVKHIALTLFNCILNCEVAFISFSSILNL 799
OY 473 AKSVLLVPLPACLNPLLYLFNPHFRDRLRPRRA-----GDSGFLAYAA 521
Db 800 VKFTILVPLPSCINPLLYTVFNHEDMGSLGKHTRFMRKSHASLISINDSVK 859
OY 528 SSCSDSTQALVAFSDVDL--LEASGRP--PGLTYGFPVTLISC 570
Db 860 RCCESTQALVFTSSITVDLPSSVSPSPAPVPSCHLSVAFYPC 906

RESULT 3

JC4301
follicle-stimulating hormone receptor
N/Alternate names: follicle-stimulating hormone receptor
C/Species: Sus scrofa domestica (domestic pig)
C/Date: 16-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 21-Jan-2000
C/Accession: JC4301
R/Hermey, J.J.; Labitb-Mansais, Y.; Yerie, M.; Bozon, V.; Couture, L.; Pajot, E.; Grebe
Gene 163, 257-261, 1995
A:Title: The porcine follicle-stimulating hormone receptor: cDNA cloning, functional expression and characterization of the receptor protein
A:Reference number: JC4301; MUID:96011644; PMID:7590277
A:Accession: JC4301
A:Molecule type: mRNA
A:Residues: 1-694 <REM>
A:Cross-references: GB:L31966
A:Experimental source: ovarian granulosa cells
C/Comment: This receptor belongs to the family of the G-protein coupled receptors. It
C/Genetics:
A:Map position: 3 q2.2-q2.3
C/Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat
C/Keywords: G protein-coupled receptor; hormone receptor; transmembrane protein
F:1-365/Domain: follicle-stimulating hormone binding #status predicted <HOB>
F:70-94/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>
F:366-388/Domain: transmembrane #status predicted <TM1>
F:398-420/Domain: transmembrane #status predicted <TM2>
F:443-464/Domain: transmembrane #status predicted <TM3>
F:485-507/Domain: transmembrane #status predicted <TM4>
F:528-549/Domain: transmembrane #status predicted <TM5>
F:573-596/Domain: transmembrane #status predicted <TM6>
F:608-629/Domain: transmembrane #status predicted <TM7>

Query Match Best Local Similarity 13.08; Score 429; DB 2; Length 694;
Matches 135; Conservative 83; Mismatches 202; Indels 104; Gaps 13;

OY 55 LEXDFPSQSSLOALDLSNNAIRSIHPEAFSTHSYKIDLDNOLTTPLAGLGLMH 113
Db 159 VERNSEFGLTFESMILNLSNGIRHINCAFNQULDELMSDNDELEPNDFHAGSG 218
OY 91 -VKLDLTDNOLTTPLAGLGLMHILKGNLALSOAFSKDSFPRKILEVYAYQCCPYG 149
Db 219 PVLDISRTIRHSLPSYGLNKLKLRARSTYVNLKLTPLKFTVLMASLYPSHCAF- 277
OY 150 MCASFKAAGOWEADLHLDDEESKRRLGLARQ 184
Db 278 -AMNRQISDLHPICNKSILRQEVDMTQARGVSLAEDGESS-----LAKE 324
OY 185 ANKHNDODLDELQLEKEDSKRHPVSQCFPTGPRPCYLFESMGIRLAVMAIVLLSVLC 244
Db 325 FDMYSEFNIDLCNEVVD-----VICSPKPDAPNFCEDINGHDIRLVIMFISLAIIG 378
OY 245 NGVLLTVFAGGAPLPVVKFVGAIAAGANTLTGICGLASVADLTGFGFSEYGARWET 304
Db 379 N-IIEVLIITSOYKLTVPFELMCNLAFTDLCIGIYLLASIDHTKSGYHVAIINWQT 437
OY 305 GLGCRATGFLAVLGSEASVLLTTAA-----VQCSVSV--SCVRAVYKSPSL 349
Db 438 GAGCDAAGFEFTVPASELSTYTLATILERHNTTHAMQLOCKQVNHASIMLXG----- 492
OY 350 GSVRAVGLGALAGLAAALPLASVEYGAASPLCLPYAPPEGAPALGPTVALVMMNSFC 409
Db 493 -----WTFATVVALEPIFGISSYKVSICLPM--DIDSPQLSVVSLVLA 540
OY 410 FVAVAGATIKLYCDLPRGDFEAVWDCAMVRHVAAILFADGLLYCYVALISFASMLGLFP 468
Db 541 FVVICGCIHITLVANPNIMSSSDTKAKKAMALITFDLCWVPSIFALISLKVPL 600
OY 469 VPPEAVKSVLLVPLPACLNPLLYLFNPHFRDRLRPRRA-----GDSGFLAYAA 521
Db 601 ITVSKLILVLFYPIVNSCANPFLVAIFTKNRRDVFILLSKFG 644

C/Accession: A41344
 R:Loosefield, H.; Mistrati, M.; Atger, M.; Salese, R.; Vu Hai-Luu Thi, M.T.; Jolivet, A.;
 Science 245, 525-528, 1989
 A/Title: Cloning and sequencing of porcine LH-hCG receptor cDNA: variants lacking transmembrane domain
 A/Accession: A41344
 A/Reference number: A41344; MUID:89332517; PMID:2502844
 A/Molecule type: mRNA
 A/Residues: 1-696 <L00>
 A/Cross-references: GB:M29525; NID:9164528; PID:AAA31062.1; PID:9164529
 C/Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat
 C/Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; membrane protein
 F:1-27/Domain: signal sequence; status predicted <SIG>
 F:28-696/Product: lutropin-choriogonadotropin hormone receptor; status predicted <LRR>
 F:99-123/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 12.7%; Score 419; DB 2; Length 696;
 Best Local Similarity 25.2%; Pred. No. 1e-24;
 Matches 134; Conservative 96; Mismatches 244; Indels 58; Gaps 12;

```

QY 26 VERSGSLPAPASLAALAAWMTASGKLEDTESOLSS-----LQALDLSMNA-IR 77
DB 153 LEIDNMHTTVPAN-AFOGMNNESTLKLNGEETIOHAFNGTTLISLEKENMAHLK 211
QY -78 SIHEAFSTLSLVKIDLDNQTITPLAGLGLMLKGNLALSOAFKSDSPKRLIL 137
DB 212 KMHDAFRGAPSIIDISTKLQALPSYGLSITQTLIATSSYLKLPSEKFTNLIDA 271
QY 138 EEPYAYQCCPY-----GMCASFKAAGQWEADLHDESSKRPGLILARQENHY 189
DB 272 TLITPHSCAFNLPTRKQNFSTIFK-----NSKQCESTARRP-----NNETLY 317
QY 190 -----DQDLDELQLEMEDSKPHPSVOCSPTRPGPKCEYLFEESGIRLAWMAVILSYLC 244
DB 318 SAIRFESSELSDMDYDYGCPSP-KTLOCAPEPDANFCEIDMGYDFKLVLMINILATIG 376
QY 245 NGVLLTVFAGGPARLPVPEVGAATGANTLTGICGLASVDALTROGSEYARRET 304
DB 377 NVTLFVLTSHYKLYP-RFLMKNLSFADFCMGYLLILASVDQVQGYHNAIDQT 435
QY 305 GIGCRATGFLAVLGESAVLLTLTAVQCSVSCVRAVGSFSGSVRAVGLCLAG 364
DB 436 GNGCSAVGFTYVASELSYTLTTLERWHITTYAIDLDKRLRIAPIMLGWLFST 495
QY 365 LAALPLASVGEYASPLCLPYAPPEGPALAGFTVALVMNNSCFVLVAGVATIKYCDL 424
DB 496 LIMPDLVGVSSYMKVSICLPM-DVETTLISQVYTLTLLINVAFTIICACIKITFAV 553
QY 425 PGGDEEAV-WDCAMVRHVAMLIFADGLLYCPVAFLSFASMGLEPTPEAVSVLLVLP 483
DB 554 ONPELMATNKDTKAKKAAVILFTDFCMAPISFPAISALAKPLITVYNSVLLVLEYP 613
QY 484 LPACINPLLYLTFNPHRDRLRLRPAGDSGLAVAAAGELKSSCDSTQA 535
DB 614 VNSCANPLLYALFTKAFRRDFELL-----LSKSGCKCKHQA 648

```

RESULT 7

JC1493

Follicle-stimulating hormone receptor - sheep

N/Alternate names: follicle stimulating hormone receptor

C/Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C/Date: 03-Feb-1994 #sequence, revision 03-Feb-1994 #text, change 21-Jan-2000

C/Accession: JCI493; MUID:93176195; PMID:8439338

R:Khan, H.; Yarney, T.A.; Saltram, M.R.

A/Title: Cloning of alternatively spliced mRNA transcripts coding for variants of ovine tg

A/Accession: JCI493; MUID:93176195; PMID:8439338

A/Molecule type: mRNA

A/Residues: 1-695 <KHA>

A/Cross-references: GB:M29525; NID:9164528; PID:AAA31062.1; PID:9164529

C/Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat

C/Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; membrane protein

F:1-27/Domain: signal sequence; status predicted <SIG>

A/Title: Molecular cloning and expression of the ovine testicular follicle stimulating hormone receptor
 A/Reference number: 147080; MUID:93351750; PMID:8994255
 A/Accession: JCI493; MUID:93351750; PMID:8994255
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-695 <VAR>
 A/Cross-references: GB:L07302; NID:9165884; PID:AAA31525.1; PID:9165885
 C/Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat
 C/Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
 F:1-27/Domain: signal sequence; status predicted <SIG>
 F:28-696/Product: lutropin-choriogonadotropin hormone receptor; status predicted <LRR>
 F:99-123/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 12.7%; Score 418.5; DB 2; Length 695;
 Best Local Similarity 26.3%; Pred. No. 1e-24;
 Matches 139; Conservative 80; Mismatches 225; Indels 85; Gaps 13;

```

QY 55 LKXDFEOLSS-SLQALDLSMAISHPAP-----STLHSL----- 90
DB 160 VERNFPMGLSPESMIVLWLSKNGIOEHNCAFNQTLDELNLSDNSNLELPNDVFOGASG 219
QY 91 -VKIDLDNQTITPLAGLGLMLKGNLALSOAFKSDSPKRLILEVYAYQCCPYG 149
DB 220 EVLIDISRTIRSLPSYGLLENLKLAKSTYHLKLPSEKFTVLEASLTYPSPHCAF- 278
QY 150 MCASFKAAGQWEADLH-----LDESSKRPGLILARQENHYDDQDLDELQ 198
DB 279 -----AMNRQTSIDLHPICKSLTROEDVDQMTQARQISLAEDDEPXYAGFDMMIS 331
QY 199 EMEDS--KPHPSVOCSPTRPGPKCEYLFEESGIRLAWMAVILSYLCNGLVLTYPAG 256
DB 332 EEDYDLCSEVVDYDYGCPSP-KTLOCAPEPDANFCEIDMGYDFKLVLMINILATIG 376
QY 257 PAPLPVPEVGAATGANTLTGICGLASVDALTROGSEYARRETSGCRATGFLAV 316
DB 391 QYKLVPRFLMKNLSFADFCMGYLLILASVDQVQGYHNAIDQT 435
QY 317 LGSEAVLLTLTAVQCSVSCVRAVGSFSGSVRAVGLCLAG 364
DB 436 GNGCSAVGFTYVASELSYTLTTLERWHITTYAIDLDKRLRIAPIMLGWLFST 495
QY 431 FASELSYTLTTLERWHITTYAIDLDKRLRIAPIMLGWLFST 495
DB 451 FASELSYTLTTLERWHITTYAIDLDKRLRIAPIMLGWLFST 495
QY 368 ALPLASVGEYASPLCLPYAPPEGPALAGFTVALVMNNSCFVLVAGVATIKYCDL 424
DB 502 LFPFGLISSYMKVSICLPM-DVETTLISQVYTLTLLINVAFTIICACIKITFAV 553
QY 428 DF-EAVWDCAMVRHVAMLIFADGLLYCPVAFLSFASMGLEPTPEAVSVLLVLP 483
DB 560 NITSSSSDRTIKKAKKAAVILFTDFCMAPISFPAISALAKPLITVYNSVLLVLEYP 613
QY 487 CLNPLLYLTFNPHRDRLRLRPAGDSGLAVAAAGELKSSCDSTQA 535
DB 620 CANPLLYALFTKAFRRDFELL-----LSKSGCKCKHQA 648

```

RESULT 8

A34548

Follicle-stimulating hormone precursor - rat

C/Species: Rattus norvegicus (Norway rat)

C/Date: 22-Jan-1993 #sequence, revision 22-Jan-1993 #text, change 13-Aug-1999

C/Accession: A34548; MUID:93351750; PMID:8994255

R:Spiegel, R.; Braun, T.; Nikolic, K.; Segaloff, D.L.; Seeburg, P.H.

A/Title: The testicular receptor for follicle stimulating hormone: structure and func

A/Accession: A34548; MUID:93351750; PMID:8994255

A/Molecule type: mRNA

A/Residues: 1-692 <SPR>

A/Cross-references: GB:L02842; NID:9204183; PID:AAA41175.1; PID:9204184

C/Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat

C/Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; transmembrane protein

F:1-27/Domain: signal sequence; status predicted <SIG>

A:Accession: JN0122
 A:Molecule type: mRNA
 A:Residues: 1-111, 'T', 113-196, 'AV', 199-306, 'A', 308-695 <MIN>
 A:Cross-references: EMBL:565085; NID:9182770; PID:AAA52477.1; PID:9182771
 C:Genetics:
 A:Gene: GDB:FSHR
 A:Cross-references: GDB:127510; OMIM:136435
 A:Map position: 2p21-2p16
 A:Introns: 223/3
 A>Note: the exact position of the intron cannot be determined from the experimental data
 C:Function:
 A:Description: receptor that mediates the biochemical effects of follitropin
 C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
 C:Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone rece
 F:1-695/Product: follitropin receptor precursor, long splice form #status predicted <SPI
 F:1-223,286-695/Product: follitropin receptor precursor, short splice form #status predi
 F:1-15/Domain: signal sequence #status predicted <SIG>
 F:16-695/Product: follitropin receptor #status predicted <SIG>
 F:16-366/Domain: extracellular hormone binding #status predicted <MAT>
 F:56-70/Domain: leucine-rich alpha-2-glycoprotein repeat #status predicted <EHB>
 F:71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
 F:96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
 F:121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F:146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
 F:172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
 F:194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
 F:367-387/Domain: transmembrane #status predicted <TM1>
 F:398-421/Domain: transmembrane #status predicted <TM2>
 F:444-465/Domain: transmembrane #status predicted <TM3>
 F:486-508/Domain: transmembrane #status predicted <TM4>
 F:529-550/Domain: transmembrane #status predicted <TM5>
 F:574-597/Domain: transmembrane #status predicted <TM6>
 F:609-630/Domain: transmembrane #status predicted <TM7>
 F:191,199,293,318/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:555/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
 F:596/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
 Query Match 12.5% Score 414.5, DB 1; Length 695;
 Best Local Similarity 25.0%; Pred. No. 2,2e-24;
 Matches 131; Conservative 94; Mismatches 233; Indels 67; Gaps 13;
 55 LKNDTFSQLS-STQALDLSWNAIRSIHPEAF-----STLNSL----- 90
 Db 160 IERNSEFVGLSFESVILMLNKNGIOEIHNCAGFTQDLDELNSDNNLEELPNDVFGASG 219
 Qy 91 -VKLDLTNOULTPLPLAGIGLMHKLKGNALASQAFSKDSEPKRILEVYAAVQCCPYG 149
 Db 220 PVILDISRTIRHSIPSYLENLKLRLARSTYNNLKLPLEKTVLMEASLTYPHCCAF- 278
 Qy 150 MCASFRRASGWEADLH-----LDDESSRRPLGLLARAQENHYDDDLDELQ 198
 Db 279 -----AMNRROISELHPICNKSILRQEDVYMOTGQSSLAEDNESSYSRGFDMTYT 331
 Qy 199 EMEDSKPPH--SVCCSPFPKPCFELFESWGIKRLAVMAIVLSVLCNGLVLTTFVAG 256
 Db 332 EFYDLCNEVVDVTCSPKPAFNPCEDILGYNILRVLIWPISTALIGNIIVLT- 390
 Qy 257 PAPIPVKPVVGAIGAGNTLTGISCGLASVDALTGQFSEYARWETGLGCRANGELAV 316
 Db 391 QYKLTVPFRFLMCNLAFADLCIGIYLLIASVDLHTSQHNVAIDMOTGAGDAGGFFTV 450
 Qy 317 LGSEASVLLTLTAAVO-----CSVSVGVRAVAGKSPSLGVSAGVAGLGLAAGLAAALP 370
 Db 451 FASELSVYTLTATLERMTTTHAQDLDC-----KVQLRHAASVVMGM-FAFAAALFP 504
 Qy 371 LASVGEYGAAPLCLPYAPPEGOPALGFTVALVMNSFCFLVNVGAYIKLYCDLPRGDF- 429
 Db 505 ITCISSTYKAVSTICLPM--DIDSPISQLYVMSLVNLVLAFAVVICGYHITLYVRNPNIV 562
 Qy 430 EAVWDCAVYRHAAMLTIFADGLLYCGVAFSLFASMLGEPVTPPEAVKSVLLVLDLPACLN 489
 Db 563 SSSSDTRIARKRAMLTIFDPLCMAPISFPALISAKLVPLITVSKAKILLVLFPHINSCAN 622

A:Accession: JN0898
 A:Molecule type: mRNA
 A:Residues: 1-695 <GRO>
 A:Cross-references: EMBL:X74454; NID:9396801; PIDN:CAA52463.1; PID:9396802
 A>Note: the authors translated the codon AGT for residue 488 as Arg
 C:Function:
 A:Description: receptor that mediates the biochemical effects of follitropin
 C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repa
 C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; phosphoprotei
 F:1-695/Product: follitropin receptor #status predicted <SIG>
 F:18-695/Product: follitropin receptor #status predicted <SIG>
 F:71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
 F:96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
 F:121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F:146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
 F:172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
 F:194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
 F:367-387/Domain: transmembrane #status predicted <TM1>
 F:399-421/Domain: transmembrane #status predicted <TM2>
 F:444-465/Domain: transmembrane #status predicted <TM3>
 F:486-508/Domain: transmembrane #status predicted <TM4>
 F:529-550/Domain: transmembrane #status predicted <TM5>
 F:574-597/Domain: transmembrane #status predicted <TM6>
 F:609-630/Domain: transmembrane #status predicted <TM7>
 F:191,199,293,318/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:555/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
 F:596/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
 Query Match 12.5% Score 414.5, DB 1; Length 695;
 Best Local Similarity 25.3%; Pred. No. 2,2e-24;
 Matches 134; Conservative 89; Mismatches 228; Indels 79; Gaps 13;
 55 LKNDTFSQLS-STQALDLSWNAIRSIHPEAF-----STLNSL----- 90
 Db 160 IERNSEFVGLSFESVILMLNKNGIOEIHNCAGFTQDLDELNSDNNLEELPNDVFGASG 219
 Qy 91 -VKLDLTNOULTPLPLAGIGLMHKLKGNALASQAFSKDSEPKRILEVYAAVQCCPYG 149
 Db 220 PVILDISRTIRHSIPSYLENLKLRLARSTYNNLKLPLEKTVLMEASLTYPHCCAF- 278
 Qy 150 MCASFRRASGWEADLH-----LDDESSRRPLGLLARAQENHYDDDLDELQ 198
 Db 279 -----AMNRROISELHPICNKSILRQEDVYMOTGQSSLAEDNESSYSRGFDMTYT 331
 Qy 199 EMEDSKPPH--SVCCSPFPKPCFELFESWGIKRLAVMAIVLSVLCNGLVLTTFVAG 256
 Db 332 EFYDLCNEVVDVTCSPKPAFNPCEDILGYNILRVLIWPISTALIGNIIVLT- 390
 Qy 257 PAPIPVKPVVGAIGAGNTLTGISCGLASVDALTGQFSEYARWETGLGCRANGELAV 316
 Db 391 QYKLTVPFRFLMCNLAFADLCIGIYLLIASVDLHTSQHNVAIDMOTGAGDAGGFFTV 450
 Qy 317 LGSEASVLLTLTAAVO-----CSVSVGVRAVAGKSPSLGVSAGVAGLGLAAGLAAALP 370
 Db 451 FASELSVYTLTATLERMTTTHAQDLDC-----KVQLRHAASVVMGM-FAFAAALFP 504
 Qy 371 LASVGEYGAAPLCLPYAPPEGOPALGFTVALVMNSFCFLVNVGAYIKLYCDLPRGDF- 429
 Db 505 ITCISSTYKAVSTICLPM--DIDSPISQLYVMSLVNLVLAFAVVICGYHITLYVRNPNIV 562
 Qy 430 EAVWDCAVYRHAAMLTIFADGLLYCGVAFSLFASMLGEPVTPPEAVKSVLLVLDLPACLN 489
 Db 563 SSSSDTRIARKRAMLTIFDPLCMAPISFPALISAKLVPLITVSKAKILLVLFPHINSCAN 622

A:Accession: JN0898
 A:Molecule type: mRNA
 A:Residues: 1-695 <GRO>
 A:Cross-references: EMBL:X74454; NID:9396801; PIDN:CAA52463.1; PID:9396802
 A>Note: the authors translated the codon AGT for residue 488 as Arg
 C:Function:
 A:Description: receptor that mediates the biochemical effects of follitropin
 C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repa
 C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; phosphoprotei
 F:1-695/Product: follitropin receptor #status predicted <SIG>
 F:18-695/Product: follitropin receptor #status predicted <SIG>
 F:71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
 F:96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
 F:121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F:146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
 F:172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
 F:194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
 F:367-387/Domain: transmembrane #status predicted <TM1>
 F:399-421/Domain: transmembrane #status predicted <TM2>
 F:444-465/Domain: transmembrane #status predicted <TM3>
 F:486-508/Domain: transmembrane #status predicted <TM4>
 F:529-550/Domain: transmembrane #status predicted <TM5>
 F:574-597/Domain: transmembrane #status predicted <TM6>
 F:609-630/Domain: transmembrane #status predicted <TM7>
 F:191,199,293,318/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:555/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
 F:596/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
 Query Match 12.5% Score 414.5, DB 1; Length 695;
 Best Local Similarity 25.3%; Pred. No. 2,2e-24;
 Matches 134; Conservative 89; Mismatches 228; Indels 79; Gaps 13;
 55 LKNDTFSQLS-STQALDLSWNAIRSIHPEAF-----STLNSL----- 90
 Db 160 IERNSEFVGLSFESVILMLNKNGIOEIHNCAGFTQDLDELNSDNNLEELPNDVFGASG 219
 Qy 91 -VKLDLTNOULTPLPLAGIGLMHKLKGNALASQAFSKDSEPKRILEVYAAVQCCPYG 149
 Db 220 PVILDISRTIRHSIPSYLENLKLRLARSTYNNLKLPLEKTVLMEASLTYPHCCAF- 278
 Qy 150 MCASFRRASGWEADLH-----LDDESSRRPLGLLARAQENHYDDDLDELQ 198
 Db 279 -----AMNRROISELHPICNKSILRQEDVYMOTGQSSLAEDNESSYSRGFDMTYT 331
 Qy 199 EMEDSKPPH--SVCCSPFPKPCFELFESWGIKRLAVMAIVLSVLCNGLVLTTFVAG 256
 Db 332 EFYDLCNEVVDVTCSPKPAFNPCEDILGYNILRVLIWPISTALIGNIIVLT- 390
 Qy 257 PAPIPVKPVVGAIGAGNTLTGISCGLASVDALTGQFSEYARWETGLGCRANGELAV 316
 Db 391 QYKLTVPFRFLMCNLAFADLCIGIYLLIASVDLHTSQHNVAIDMOTGAGDAGGFFTV 450
 Qy 317 LGSEASVLLTLTAAVO-----CSVSVGVRAVAGKSPSLGVSAGVAGLGLAAGLAAALP 370
 Db 451 FASELSVYTLTATLERMTTTHAQDLDC-----KVQLRHAASVVMGM-FAFAAALFP 504
 Qy 371 LASVGEYGAAPLCLPYAPPEGOPALGFTVALVMNSFCFLVNVGAYIKLYCDLPRGDF- 429
 Db 505 ITCISSTYKAVSTICLPM--DIDSPISQLYVMSLVNLVLAFAVVICGYHITLYVRNPNIV 562
 Qy 430 EAVWDCAVYRHAAMLTIFADGLLYCGVAFSLFASMLGEPVTPPEAVKSVLLVLDLPACLN 489
 Db 563 SSSSDTRIARKRAMLTIFDPLCMAPISFPALISAKLVPLITVSKAKILLVLFPHINSCAN 622

Db 499 AALPFIPIGSIYMKVYSICLPM--DIDSPLSQLYVMSLLVLNLAVLVCIGCTTHLYLV 556
 QY 425 PRDGF-EAWWDCAMRHVAMLIIFADGLLCPVAFLESPASMLGFPPTPAVKSVLLVLP 483
 Db 557 RNNIVYSSSDTIRAKMAMLIITDFLCMAPISEFPAISAKLPILITVSKAKILLVLP 616
 QY 484 LPPCLINLILFLFPHRDLRLPRAGDSGPLAAVAAAGELEKSCDST 533
 Db 617 INSCANFLIATFTKNRRDFILLSKFG-----CYEMQOIVRTETST 661

RESULT 11

thyrotropin receptor precursor - dog
 A:Accession: A40077
 N:Alternate names: thyroid-stimulating hormone receptor; TSH receptor
 C:Species: Canis lupus familiaris (dog)
 C>Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 13-Aug-1999
 C:Accession: A40077; S06933
 R:Parmentier, M.; Libert, F.; Maenhaut, C.; Lefort, A.; Gerard, C.; Perret, J.; Van Sande
 Science 246, 1620-1622, 1989
 A>Title: Molecular cloning of the thyrotropin receptor.
 A:Reference number: A40077; MUID:90084524; PMID:2556796
 A:Accession: A40077
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-764 <PAR>
 A:Cross-references: GB:M29957; NID:g164098; PIDN:AAA30901.1; PID:g164099
 R:Parmentier, M.; Libert, F.; Maenhaut, C.; Lefort, A.; Gerard, C.; Perret, J.; van Sande
 Nucleic Acids Res. 17, 10493, 1989
 A>Title: Nucleotide sequence of the dog thyrotropin receptor cDNA.
 A:Reference number: S06933; MUID:90098886; PMID:2602159
 A:Accession: S06933
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-764 <PAR>
 A:Cross-references: EMBL:X17146; NID:9849; PIDN:CAA5026.1; PID:9850
 C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat
 C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-764/Product: thyrotropin receptor #status predicted <MAT>
 F:53-76/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LR1>
 F:77-101/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:102-126/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LR3>
 F:127-151/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LR4>
 F:152-176/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LR5>
 F:179-200/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LR6>
 F:201-225/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LR7>

Query Match 12.5%; Score 414.5; DB 2; Length 764;
 Best Local Similarity 23.7%; Pred. No. 2.5e-24;
 Matches 141; Conservative 102; Mismatches 258; Indels 95; Gaps 12;

QY 33 SLPAHPASTAALASNTTASGKLENDTFSQ-----SLQALDLSWNA-IRSHPEAF 84
 Db 166 STPAN-----AFQGLCNFTLTLYKLNNGFTSIQGHAFNGTLDVAVLNKKNYLAIIDKDAF 221
 QY 85 STLHS-LVKLIDLTDNQLTTLPLAGLGLMHLKLGKLNLAISQAFSKDSFPKRLILEVPYAY 143
 Db 222 GGVYSGPTLDDISYTSVYALPSPKGLEHLKELIARNMTWLKRLPLSLFHLTRADLSYPS 281
 QY 144 QCCPY-----GMCASF-----FKASG 159
 Db 282 HCCAFKNOKKIRGITLESIMCNESIRSRLORKSVNTLNGPPDEYEYELGDSHAGYDNS 341
 QY 160 QW-----EAEDLHLDDESSKRPGLGLARQAEHNYDQDLDELQLEMEDSKP 205
 Db 342 QPQDTHSNHYVYFEEQDEILIGFQGLKKNPQETLQAFDNHYDYTV-----CGG 392
 QY 206 HPSVQCSPTPEGPKPCPEYLEFSMGIRLAWAIVLISVLCNGLVLTIVFAGGAPAPLPVKF 265
 Db 393 NEDWCTPKSDEFPNCEDIMGYKFLRIYVWFVSLALIGNVFLIVLTISHKLTVP-RF 451

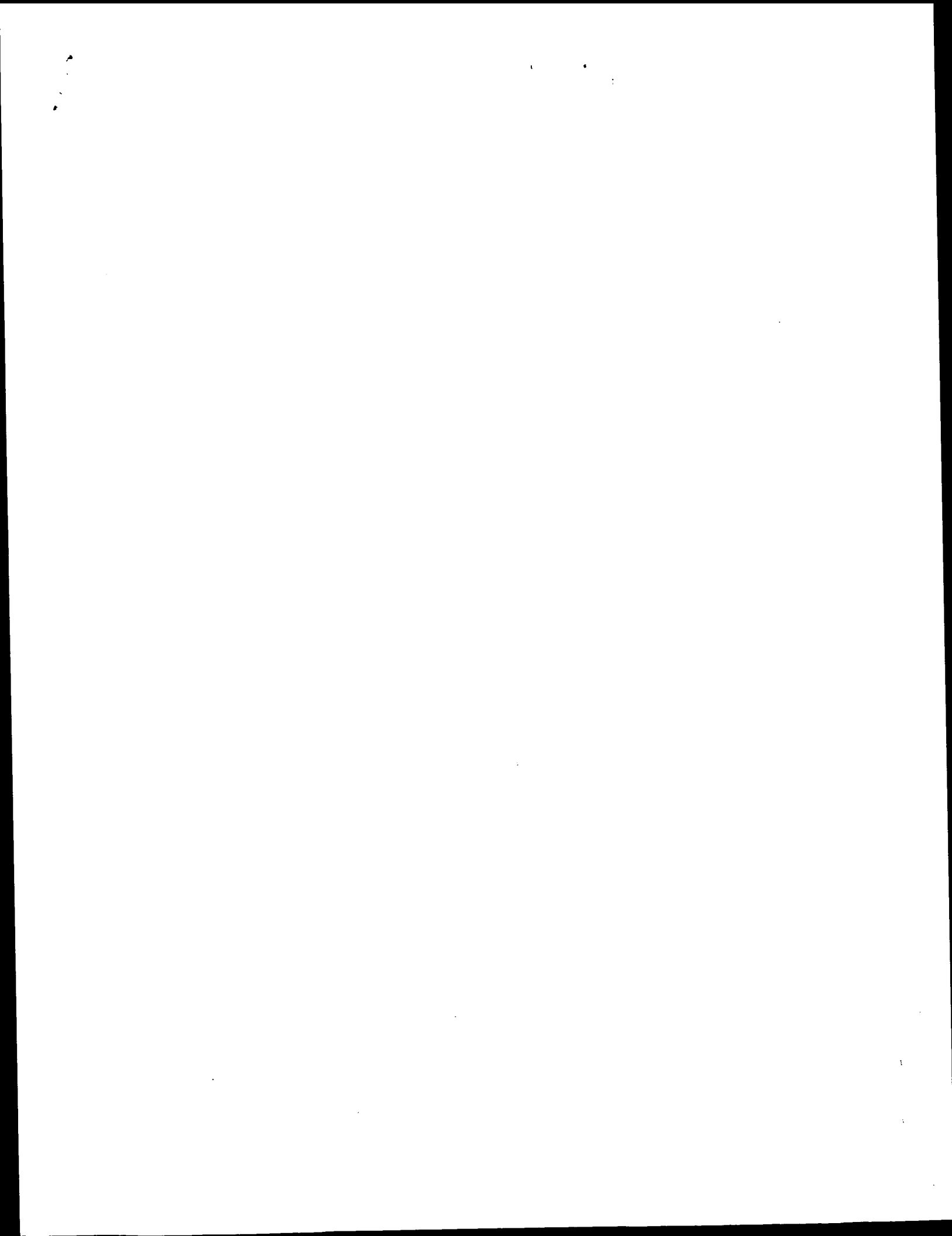
QY 266 VVGAIAGANTLTIGISGGLASVADLTFFGQSEYGARMETGICRATGFLAVLGSEASVL 325
 Db 452 LMCNLAFDFQCGMWLILLIASVLYTHSEYVNHAIQDMGTGCGCTAGFTVFASELSTYT 511
 QY 326 LTLAAQCSVSVSCYCRANGKSPSLGSRAGVLCGLALAGLAALPLASVGEYASPLCLP 385
 Db 512 LVITLIERVYATTFMRDLRKLRLHAYAVIMGWCCFLALALPLVVISSTAYKVICLP 571
 QY 386 YAPBESQALFEYVALVMNSFCFLVAVAGYIKLYCDLPRGDFE-AVWDCAMRHVAM 444
 Db 572 M-DIETPLATLIIIVLLNLVAFIVYVSCVYKITYVRNPQVNGDKDKRIAKRMAYL 629
 QY 445 IFADGLLCPVAFLESPASMLGFPPTPAVKSVLLVLPPLPCLMPLILFPHRDL 504
 Db 630 IFTDFMCMAPISFPYALSAIMNRPILITVNSKILLVLEFPLNSCANPFLAITTKAFQDV 689
 QY 505 RLRRPAGDSGPLAAVAAAG-----ELEKSCDSTQALVAFSDVLTLEAS 549
 Db 690 FILLKFGICCKROAQVIRGQVSPKNSAGIQIQIKYTRDKRQSLPMNODYELENS 745

RESULT 12

JC5643
 thyroid stimulating hormone receptor precursor - sheep
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C>Date: 28-Oct-1997 #sequence_revision 28-Oct-1997 #text_change 20-Jun-2000
 C:Accession: JC5643
 R:Bockmann, J.; Wintner, C.; Wilkowskl, W.; Kreutz, M.R.; Boeckers, T.M.
 Biochem. Biophys. Res. Commun. 238, 173-178, 1997
 A>Title: Cloning and expression of a brain-derived TSH receptor.
 A:Reference number: JC5643; MUID:97445147; PMID:9299474
 A:Accession: JC5643
 A:Molecule type: mRNA
 A:Residues: 1-764 <BOC>
 A:Cross-references: GB:Y13434; NID:92462632; PIDN:CAA73846.1; PID:92462633
 A:Experimental source: hypothalamus
 A>Note: the sequences of residues 4-7 and 8-17, 4136-439 and 440-449 are interchanged
 C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat
 F:1-21/Domains: signal sequence #status predicted <SIG>
 F:22-764/Product: thyroid stimulating hormone receptor #status predicted <MAT>
 F:179-200/Domains: extracellular #status predicted <EXC>
 F:303-382/Domains: transmembrane #status predicted <TM1>
 F:416-439/Domains: transmembrane #status predicted <TM2>
 F:451-474/Domains: transmembrane #status predicted <TM3>
 F:495-518/Domains: transmembrane #status predicted <TM4>
 F:538-561/Domains: transmembrane #status predicted <TM5>
 F:581-604/Domains: transmembrane #status predicted <TM6>
 F:627-650/Domains: transmembrane #status predicted <TM7>
 F:661-684/Domains: transmembrane #status predicted <TM8>

Query Match 12.5%; Score 414.5; DB 2; Length 764;
 Best Local Similarity 25.0%; Pred. No. 2.5e-24;
 Matches 143; Conservative 93; Mismatches 244; Indels 93; Gaps 15;

QY 33 SLPAHPASTAALASNTTASGKLENDTFSQ-----SLQALDLSWNA-IRSHPEAF 84
 Db 166 SVAN-----AFQGLSNFTLTLYKLNNGFTSIQGHAFNGKIDVAVLNKKNYITVIDQAF 221
 QY 85 STLHS-LVKLIDLTDNQLTTLPLAGLGLMHLKLGKLNLAISQAFSKDSFPKRLILEVPYAY 143
 Db 222 AGVYSGPTLDDISYTSVYALPSPKGLEHLKELIARNMTWLKRLPLSLFHLTRADLSYPS 281
 QY 144 QCCPY-----GMCASF-KASGQW----- 161
 Db 282 HCCAFKNOKKIRGITLQSLMCNNESSIMGLRORSASALNCPYOEYEDLIGDSAGYKENS 341
 QY 162 EADDLH-----LDDDESS-----KRPGLARQAEHNYDQDLDELQLEMEDSKP 205
 Db 342 KFDQTHSNHYVYFEEQDEILIGFQGLKKNPQETLQAFDNHYDYTVGCGSEEM----- 396
 QY 206 HPSVQCSPTPEGPKPCPEYLEFSMGIRLAWAIVLISVLCNGLVLTIVFAGGAPAPLPVKF 265



223 LTVWFIEFLVALFENLLVILTFASCTS-LPSSKLEIGLISVSNLFMGITYTGILTFELDAVS 281

[illegible]

```

RESULT 2
US-09-862-767A-4
; Sequence 4, Application US/09862767A
; Patent No. US20020034786A1
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: TANGO-72 AND TANGO-77 NUCLEIC ACID MOLECULES AND POLYPEPTIDES
; FILE REFERENCE: MI011997-008P1RCP1CN1(W)
; CURRENT APPLICATION NUMBER: US/09/862,767A
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/127,856
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 60/054,646
; PRIOR FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Homo sapiens
; SS-09-862-767A-4

```

[illegible]

```

Db      281  SSSSIKIKVAMIIIFNCFPCVPAAFEPSAPLITAIISPEIMKSVTLIEPLPLACLNPLY 340
Oy      493  YLLFMPHPRDD-----LRRLRAPDGSGFLAAVAAAGLEK-----SSCDS 532
Db      341  YVFEPKREKEMKLLKRRVTKSKSSVSYSISQGGCLFQDPRYYDGGMYSHQGNLTVDC 400
Oy      553  TQALVAFSDVD---LILEASAGRPRLGLEYGPSVTLISQOP 573
Db      401  CESFLTKPVSCKHLL-----KSHSCALPLVAASQRP 432

```

```

RESULT 3
US-09-862-767A-8
: Sequence 8, Application US/09862767A
: Patent No. US20020034786A1
: GENERAL INFORMATION:
: APPLICANT: Pan, Yang
: TITLE OF INVENTION: TANGO-72 AND TANGO-77 NUCLEIC ACID MOLECULES AND POLYPEPTIDES
: FILE REFERENCE: MBIO1997-00861RCP1CN1(M)
: CURRENT APPLICATION NUMBER: US/09/862,767A
: CURRENT FILING DATE: 2001-05-21
: PRIOR APPLICATION NUMBER: US 09/127,856
: PRIOR FILING DATE: 1998-08-03
: PRIOR APPLICATION NUMBER: US 60/054,646
: PRIOR FILING DATE: 1997-08-04
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 8
: LENGTH: 431
: TYPE: PR1
: ORGANISM: Homo sapiens
: US-09-862-767A-8

```

Query Match	23.7%;	Score 783;	DB 10;	Length 431;
Best Local Similarity	37.4%;	Pred. No. 5.3e-54;		
Matches 173;	Conservative 73;	Mismatches 157;	Indels 60;	Gaps 9

[illegible]

RESULT 4
US-09-862-767A-6
: Sequence 6, Application US/09862767A

Patent No. US20020034786A1
GENERAL INFORMATION:
APPLICANT: Pan, Yang
TITLE OF INVENTION: TANGO-72 AND TANGO-77 NUCLEIC ACID MOLECULES AND POLYPEPTIDES
FILE REFERENCE: MB101997-008P1C1C1N1(M)
CURRENT APPLICATION NUMBER: US/09/862,767A
CURRENT FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: US 09/127,856
PRIOR FILING DATE: 1998-08-03
PRIOR APPLICATION NUMBER: US 60/054,646
PRIOR FILING DATE: 1997-08-04
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 6
LENGTH: 358
TYPE: PRT
ORGANISM: Homo sapiens
US-09-862-767A-6
Query Match 22.0% Score 727.5; DB 10; Length 358;
Best Local Similarity 47.0%; Pred. No. 9.3e-50; Indels 3; Gaps 2;
Matches 142; Conservative 54; Mismatches 103; Indels 3; Gaps 2;
209 VCCSPGPKFCEYLFESEWIRLAVNAIVLISVLCNGVLTTFVAGGAPDLPPVKYVG 268
29 IHCPTGAFKPEYILGSMIRLTVETFLVALFENLVLITTFASCTSLPSKLFIG 87
269 AIGANTLIGISGLASVDALTFGQSEYGARWETGLCRATGFLAVGSEASVLLTL 328
88 LISNLFNMTYIGILFLDVAWSGRFAEIGWETGSCYAGFLAVSSSAIFLLML 147
329 AAVCCSVSCVARYGSPSLGSRVAGVLCGLALAGLAAALPLASGEAGAPLCPYAP 388
148 ATVERISAKDIMKNGSNHLKOFERVAALLAFGATVAGCFPLHREXESASPLCLPF 205
389 PEGOPALGFTVALVMNSFCFLVAVAGYIKYCDLPRGDFEAWMOCAMVAVAMLIAD 448
206 PIGEPISLGFYTVLILNLSIAFLAMAVLYTKLYCNLEKEDLSNOSSMIKHVMILFTN 265
449 GLYCPVAFSLASMLGLFPVPEAVKSVLLVLLPLACLNPLIYLLFNPHPRDLRLR 508
266 CIEFCVAFSFAPLTALSISPELMSVTLIFPLPACLNPLVLYVFNPKFKEDWMLK 325
QY 509 PR 510
DB 326 RR 327
RESULT 5
US-09-804-551B-20
Sequence 20, Application US/09804551B
Patent No. US20020056151A1
GENERAL INFORMATION:
APPLICANT: Bayer Aktiengesellschaft
TITLE OF INVENTION: Receptors for peptides from insects
FILE REFERENCE: Le A 34 394
CURRENT APPLICATION NUMBER: US/09/804,551B
CURRENT FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: DE 100 13 618.4
PRIOR FILING DATE: 2000-03-18
NUMBER OF SEQ ID NOS: 92
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 20
LENGTH: 861
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-804-551B-20
Query Match 16.3%; Score 539; DB 10; Length 861;
Best Local Similarity 27.7%; Pred. No. 1.7e-34; Indels 126; Gaps 22;
Matches 174; Conservative 95; Mismatches 234; Indels 126; Gaps 22;
QY 33 SLPAHPSIALAASNTTASGLKLEXTFSQL-SSLQALDLSMNAIRSIHPEAF----- 84

DB 10 SLPEGLSKLSQLOELIST-SNRLRMINDTELPSSMQMLDRAPLSTISPGAFRGSKLR 68
QY 85 -----STLS-----LVKLDLTNDQTLT- 103
DB 69 KILSDVRLRSFPELECHALEILIKDRAGIOEVANLQGRFRLKSLDLSNQIEK 128
QY 104 --PLAGLGIMLTKGNALSAFSDSF--PKRILEPVAYOCPCYCMCAFESKAS 158
DB 129 GRPFNGIKQINDLLSYNI--KALPDFAFOGIPKLOLLKOLGASMDP----- 175
QY 159 GOMEAEIDLHDESSKRLGLLARGAENHYDDDE-----LOLEMDSK 204
DB 176 --WETA--INFNEBOLQOTGGATSTMEYFEHDSVPATYGFGTGLFSGMSTEDFQ 232
QY 205 PHSVCCSPGPKFCEYLFESEWIRLAVNAIVLISVLCNGVLTTFVAGGAPDLPPVK 264
DB 233 P-GSVQCLPMGPFLPCADLFDMWTLKRGVAVVFLSLGNGTVAVLLC-SRSKMDYPR 290
QY 265 FVGAIGANTLIGISGLASVDALTFGQSEYGARWETGLCRATGFLAVGSEASV 324
DB 291 FIVCNIAADEFMGIYDILAVDAITGEFRMPALPMOMSVLCQSLGFLAVLSSELSV 349
QY 325 LTLAAVCCSVSCVARYGSPSLGSRVAGVLCGL--ALAGLAAALPLASGEAGAPL 408
DB 350 -YTLAVITLERVYATHTAHLNKRSLKQAGTMSVGVFALIMAPLVGSDYRKFVAV 439
QY 383 CLPYAPPEGOPALGFTVALVMNSFCFLVAVAGYIKYCDLPRGDFEAVV--DCAMVR 464
DB 409 CLPFTTGG--PASTLYVISLMFNGCAFLTLGVCYKMKWAL-RG--SQANNTNDSRIAK 499
QY 440 HVMLIPADGLYCPVAFSLASMLGLFPVPEAVKSVLLVLLPLACLNPLIYLLFNPH 524
DB 465 RMALVETDFELCWSPIAFSITITAFGLQILSLQAKITFVFLPLNCSNPFLYIMTKQ 549
QY 500 FRDLRLRPR-----AGDSGLAVYAAAGLEKSSQDSQALVAFSDVL--ILEASE 550
DB 525 FRKDCVTLCKHFEESRVYGGGPGRGAVARTKG-----DLRPLLPAA 570
QY 551 AGRPGLETYGFPSYTLISCCQCAPRIE 579
DB 571 VAHPGCR-----CLRMPLSEKPMNMKME 594
RESULT 6
US-09-877-804-7
Sequence 7, Application US/09877804
Patent No. US20020061557A1
GENERAL INFORMATION:
APPLICANT: Nikolic, Karoly
APPLICANT: McFarland, Keith C.
APPLICANT: Seeburg, Peter H.
TITLE OF INVENTION: Glycoprotein Hormone Receptor Molecules
FILE REFERENCE: P0576P1C2
CURRENT APPLICATION NUMBER: US/09/877,804
CURRENT FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: US 08/207,814
PRIOR FILING DATE: 1994-03-07
PRIOR APPLICATION NUMBER: US 07/781,153
PRIOR FILING DATE: 1991-10-31
PRIOR APPLICATION NUMBER: US 07/347,683
PRIOR FILING DATE: 1989-05-05
NUMBER OF SEQ ID NOS: 22
SEQ ID NO 7
LENGTH: 675
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: deduced sequence
US-09-877-804-7
Query Match 12.6%; Score 417; DB 10; Length 675;

Best Local Similarity 25.0%; Pred. No. 4.7e-25;
Matches 167; Conservative 86; Mismatches 251; Indels 164; Gaps 23;

```

QY 16 SGIPSTHASYERSOGL-----SLPAHPASIALAASNTASGKLEXTDFSSLSL 66
Db 75 SNLKLHEIRLEKANNLLYINPEAFOLP-----SLRYLLISNT---GIRKLPAVHIQSL 127
QY 67 Q-----ALDSMNAIRSIHPEAF-----S 85
Db 128 QKVLIDIDNINIHIVARNSEFMGLSFESVILWLSKNGIEIHNCFAFNGTQDELNLSDNN 187
QY 86 TLHSL-----VKLDLTDNOLTTPLAGLGLMLKLGKGNLALSQAFSKDSFPL 134
Db 188 NLEELPNDVFOGASGPVILIDISTRTVHSLPNHGLENLKRLARSTYRLKLPNDKEVT 247
QY 135 RILEVPAVYOCPCYGCASCFKASGQWEAEHL-----LD 169
Db 248 MASLTPYSHCCAF-----ANLKRQISLHPICNKSILRQIDDMTOIGDQVSLID 239
QY 170 DEESKRPLGLARQAEHNYDODLELQEMEDSKPHPSVOCSPTPGPKCEYLFESWG 229
Db 300 DEPS-----YKGSDDMYNEFDYDLCN---EYVD-----VTCSPKPAFNPCEIDINGYNI 346
QY 230 IRLAWMAIVLISVLCNGLVLLTVFAGGPAPLPVKFVVGAIAGANTLTGICGLASVDA 289
Db -347 LRVLMFISLITIGNTTVLV-VLTTSQYKLTVPFLMCLNLAFAADLCIGIYLLIASVDI 405
QY 290 LTFQGFSEYGARWETGLGCRATGFLAVLGEASVLLTLAAVQ-----CSVSVSCVRAY 343
Db 406 HTSQYHNHAYIDMOTGAGCDAGFFVFASLSVTLTATLERMTTHAMQEC-----461
QY 344 GKSPISGVSRAGVLCGLALAGLAALPLASVGEYASPLCLPYAPREGOPALGFTVALV 403
Db 462 -KVOLRHAASVWVLG-WTFAPAAALPFIIGISSYMKVSIQCPM--DIDSPLSQLYVALL 517
QY 404 MNNSCFELVAGAYIKLYCDLPRGDF-EAVWDCAMVRHVAMLIFADGLLYCPVAFISFAS 462
Db 518 VLVNLAFAVYICGCTHYIYLVNRPITVSSSDTKIAKRMATLIFTDLCAPIISFAISA 577
QY 463 MGLFPVPEAVKSVLLVLPACLNPLLYLFPNPRDLRLRRRAGDSGLAYAAA 522
Db 578 SLKVPPLITVSRAKILLVLYFPIINSCANPFLYATFTKFRDFIL-----622
QY 523 GLEKSSCDSTQALVAFSDVDLLEASAGRPGLFTYGPFSVTLISCOOPGAPRLGESH 582
Db 623 --LSKFGCEYEMAOIYRT-----ETSSA-----THNF-HARKSHCS--SAPRTNYSY 664
QY 583 CVEPEGNH 590
Db 665 VLVPPL-NH 671

```

RESULT 7
US-09-877-804-6
Sequence 6, Application US/09877804
Patent No. US20020061557A1
GENERAL INFORMATION:
APPLICANT: McFarland, Karoly
APPLICANT: Nikolic, Karoly
APPLICANT: Segalo, Deborah L.
APPLICANT: Seeburg, Peter H.
TITLE OF INVENTION: Glycoprotein Hormone Receptor Molecules
FILE REFERENCE: P0576P1C2
CURRENT APPLICATION NUMBER: US/09/877,804
PRIOR FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: US 08/207,814
PRIOR FILING DATE: 1994-03-07
PRIOR APPLICATION NUMBER: US 07/781,153
PRIOR FILING DATE: 1991-10-31
PRIOR APPLICATION NUMBER: US 07/347,683
NUMBER OF SEQ ID NOS: 1989-05-05
SEQ ID NO 6

LENGTH: 692
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: deduced sequence
US-09-877-804-6

Query Match 12.6%; Score 417; DB 10; Length 692;
Best Local Similarity 25.0%; Pred. No. 4.8e-25;
Matches 167; Conservative 86; Mismatches 251; Indels 164; Gaps 23;

```

QY 16 SGIPSTHASYERSOGL-----SLPAHPASIALAASNTASGKLEXTDFSSLSL 66
Db 92 SNLKLHEIRLEKANNLLYINPEAFOLP-----SLRYLLISNT---GIRKLPAVHIQSL 144
QY 67 Q-----ALDSMNAIRSIHPEAF-----S 85
Db 145 QKVLIDIDNINIHIVARNSEFMGLSFESVILWLSKNGIEIHNCFAFNGTQDELNLSDNN 204
QY 86 TLHSL-----VKLDLTDNOLTTPLAGLGLMLKLGKGNLALSQAFSKDSFPL 134
Db 205 NLEELPNDVFOGASGPVILIDISTRTVHSLPNHGLENLKRLARSTYRLKLPNDKEVT 264
QY 135 RILEVPAVYOCPCYGCASCFKASGQWEAEHL-----LD 169
Db 265 MASLTPYSHCCAF-----ANLKRQISLHPICNKSILRQIDDMTOIGDQVSLID 316
QY 170 DEESKRPLGLARQAEHNYDODLELQEMEDSKPHPSVOCSPTPGPKCEYLFESWG 229
Db 317 DEPS-----YKGSDDMYNEFDYDLCN---EYVD-----VTCSPKPAFNPCEIDINGYNI 363
QY 230 IRLAWMAIVLISVLCNGLVLLTVFAGGPAPLPVKFVVGAIAGANTLTGICGLASVDA 289
Db 364 LRVLMFISLITIGNTTVLV-VLTTSQYKLTVPFLMCLNLAFAADLCIGIYLLIASVDI 422
QY 290 LTFQGFSEYGARWETGLGCRATGFLAVLGEASVLLTLAAVQ-----CSVSVSCVRAY 343
Db 423 HTSQYHNHAYIDMOTGAGCDAGFFVFASLSVTLTATLERMTTHAMQEC-----478
QY 344 GKSPISGVSRAGVLCGLALAGLAALPLASVGEYASPLCLPYAPREGOPALGFTVALV 403
Db 479 -KVOLRHAASVWVLG-WTFAPAAALPFIIGISSYMKVSIQCPM--DIDSPLSQLYVALL 534
QY 404 MNNSCFELVAGAYIKLYCDLPRGDF-EAVWDCAMVRHVAMLIFADGLLYCPVAFISFAS 462
Db 535 VLVNLAFAVYICGCTHYIYLVNRPITVSSSDTKIAKRMATLIFTDLCAPIISFAISA 594
QY 463 MGLFPVPEAVKSVLLVLPACLNPLLYLFPNPRDLRLRRRAGDSGLAYAAA 522
Db 595 SLKVPPLITVSRAKILLVLYFPIINSCANPFLYATFTKFRDFIL-----639
QY 523 GLEKSSCDSTQALVAFSDVDLLEASAGRPGLFTYGPFSVTLISCOOPGAPRLGESH 582
Db 640 --LSKFGCEYEMAOIYRT-----ETSSA-----THNF-HARKSHCS--SAPRTNYSY 681
QY 583 CVEPEGNH 590
Db 682 VLVPPL-NH 688

```

RESULT 8
US-09-804-626-8
Sequence 8, Application US/09804626
Patent No. US20020128190A1
GENERAL INFORMATION:
APPLICANT: Lobel, Leslie
APPLICANT: Joyce
TITLE OF INVENTION: EXPRESSION OF PROPERLY FOLDED AND SOLUBLE EXTRACELLULAR DOMAIN
FILE REFERENCE: 0575/62259/JPM/SHS
CURRENT APPLICATION NUMBER: US/09/804,626
NUMBER OF SEQ ID NOS: 2001-03-09

```

; SOFTWARE: Patentin version 3.0
; SEQ ID NO 8
; LENGTH: 695
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-804-626-8

```

```

Query Match      12.5%; Score 414.5; DB 10; Length 695;
Best Local Similarity 25.1%; Pred. No. 7.6e-25;
Matches 123; Conservative 94; Mismatches 231; Indels 43; Gaps 11;

```

```

QY 64 SSIALDLSSWNAIRSHPEASTLSIVKLDITDNLTLPLAGLGGLMLKIKGLAL 122
DB 193 TQDAVNLSDNNLEELPNDVFGASGPVLLDSRRIRHSLPSVGLLENLKKLRARSTYNL 252
QY 123 SQAFKSDPFKRLILEVPAVYOCPCPYGMCASFRAQSCQMAEDLH-----LDDE 171
DB 253 KKLPTLEKLVALMEASLTYPSSHCCAF-----ANWRROISELHPICNLSILROEVDYM 304
QY 172 ESKRPLGLARQAEHNYDDDELQLEMEDSKRHP--SVQSPPTGPPKPCCEYLFECSWG 229
DB 305 TQARGORSSLAEDNESSYSGFDWTTTEFDYDLCNEVVDYTCSPKPAFNCEIDIMGYN 364
QY 230 IRLAVMAIVLTVLNGVLLTVFAGGPAPLPVKFVVGALAGANTLTGISCGLLASYDA 289
DB 365 LRVILWIFISILATIGN-IIVLVILTSQYKLTVPRLMCLNAFADLCIGIYLLIASYDI 423
QY 290 LTFGESEYCARWETGICRATGFLAVLGSEASVLLTLAAV-----GCVSVSCVRAV 343
DB 424 HTKQYHNYAIDMOTGAGCDAAGEFTVFASELVYTLATLRLMHTTHAMQDLC----- 479
QY 344 GKSSLSVVRAGVGCALAGLAAALPLASVGEYGASPLCLPYAPREGOPALGFTVALV 403
DB 480 -KVQDRAASVWVGWMT-FAPRAALFPIFGISYWKVASICLPM--DIDSPISQIYVMSLL 535
QY 404 MNNSFCLVAGAYIKLYCDLPRGDF-EAVWDCANVRHVAAILFADGLLXCPVALLSAS 462
DB 536 VLVNIAFVIGCYVHIHLYFRNPVIYSSSDTRIAKRAMAILFDLCMAISFALISA 595
QY 463 MGLGFPVPEAVKSVLLVPLPACLNPLLYLLFNPHFRDRLRLPRASGSLPAYAA 522
DB 596 SLKVPVLTVSAKILLVLFHPINSCANPFLVAFITKNRRDFILLSKCG-----CYEMQ 650
QY 523 GELEKSCDST 533
DB 651 AQIYRTETST 661

```

```

RESULT 9
US-09-877-804-3
; Sequence 3, Application US/09877804
; Patent No. US20020061557A1
; GENERAL INFORMATION:
; APPLICANT: Nikolic, Karoly
; APPLICANT: McFarland, Keith C.
; APPLICANT: Segalo, Deborah L.
; APPLICANT: Seeburg, Peter H.
; TITLE OF INVENTION: Glycoprotein Hormone Receptor Molecules
; FILE REFERENCE: P0576P1C2
; CURRENT APPLICATION NUMBER: US/09/877,804
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 08/207,814
; PRIOR FILING DATE: 1994-03-07
; PRIOR APPLICATION NUMBER: US 07/781,153
; PRIOR FILING DATE: 1991-10-31
; PRIOR APPLICATION NUMBER: US 07/347,683
; PRIOR FILING DATE: 1989-05-05
; NUMBER OF SEQ ID NOS: 22
; SEQ ID NO 3
; LENGTH: 674
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:

```

```

; OTHER INFORMATION: deduced sequence
US-09-877-804-3

```

```

Query Match      12.3%; Score 407.5; DB 10; Length 674;
Best Local Similarity 24.0%; Pred. No. 2.6e-24;
Matches 135; Conservative 99; Mismatches 254; Indels 75; Gaps 13;

```

```

QY 47 SNTTASGKLEXPFSQLS-----LQALDLSMNA-IRSHPEASTLSIVKLDITDN 98
DB 151 NNSSVTLKLYGNNGFEVOSHAFNGTTLISLEKENITLKHSHGAFGAGATPDLDSST 210
QY 99 QLTTLPLAGLGGLMLKGLALSQAFKSDPFKRLILEVPAVYOCPCPYGMCASFRAK 158
DB 211 KLOALPSHGLESITLALSSYSLKTPSKKFTSLVATLTYSHCCAFRNLK----- 265
QY 159 GQWEAEDLHLDDESSKRPGLLARQAEH--YDQDDELQLEMED-----SKRHPVQC 211
DB 266 ---KEQNPFSFIFENFSKQCESTVRKADNETLYSAIFEENELSGMDVDYGCSP-KTLQC 321
QY 212 SEPTGPPKPCCEYLENSGIRLAVMAIVLTVLNGVLLTVFAGGPAPLPVKFVVGAI 271
DB 322 APEPDAFNPCCEDINGYAFRLVILMLTNILALFGN-LTVLFULLISRYKLTVPRLMCL 380
QY 272 GANLTLGISCGLASVDAITFGQSEYCARWETGICRATGFLAVLGSEASVLLTLAAV 331
DB 381 FADFCMGILYLLIASVDSQTKQYVNHAIIDMOTGSCGCAAGFTVFASELVYTLVITL 440
QY 332 QCSVSVSCVRAVYKSPSLGVRAGVGLALAGLAAALPLASVGEYGASPLCLPYAPREG 391
DB 441 ERWHTITVAVOLDQKLRIRHAIPIMLGMLFSTLTATMTLVGINSYMKVSLCPM--DYE 498
QY 392 QPALGFTVALVNMNSFCLVAGAYIKLYCDLPRGDFEAV-WDCANVRHVAAILFADGL 450
DB 499 STLQSVYTLISLILLVNVAFAVVICACIIRITFAVQNPBELTAPKDKIRAKMALLFTDPT 558
QY 451 LYCPAFLSFASMLGLFPVTPAVKSVLLVPLPACLNPLLYLLFNPHFRDRLRLR 510
DB 559 CMAISFALISAFAVPLITVTNSKILLVLFYVNSCANPFLVAFITKRFQDFILLSR 618
QY 511 AGDSGPLAVAAAGELEKSCDSTQALVAFSDVLLLEASEACRPPCLETY-----GFP- 563
DB 619 FG-----CCKRA-----ELYRRKFSAYTNSCKNGFPG 647
QY 564 -----SVTLISQCGQGAPR 577
DB 648 ASKPSQATIKLSTVHQQPIPPR 670

```

```

RESULT 10
US-09-877-804-2
; Sequence 2, Application US/09877804
; Patent No. US20020061557A1
; GENERAL INFORMATION:
; APPLICANT: Nikolic, Karoly
; APPLICANT: McFarland, Keith C.
; APPLICANT: Segalo, Deborah L.
; APPLICANT: Seeburg, Peter H.
; TITLE OF INVENTION: Glycoprotein Hormone Receptor Molecules
; FILE REFERENCE: P0576P1C2
; CURRENT APPLICATION NUMBER: US/09/877,804
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 08/207,814
; PRIOR FILING DATE: 1994-03-07
; PRIOR APPLICATION NUMBER: US 07/781,153
; PRIOR FILING DATE: 1991-10-31
; PRIOR APPLICATION NUMBER: US 07/347,683
; PRIOR FILING DATE: 1989-05-05
; NUMBER OF SEQ ID NOS: 22
; SEQ ID NO 2
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:

```


QY 502 DD-----LRLAPRAGDSGLAYAAAGLEK-----SSCDSTQALVAFSD 541
 Db 119 EDWLKLRKRRYKSSGSVSISISOGGLEQDFYDQGMVSHLGNIITVDCCESSFLTKP 178
 QY 542 VD-----LILASEAGRPGLFETGPPSVTLISCOOP 573
 Db 179 VSCKHIT-----KSHSCPALAVASCORP 201

RESULT 13

US-09-862-767A-9
 ; Sequence 9, Application US/09862767A
 ; Patent No. US20020034786A1
 ; GENERAL INFORMATION:

; APPLICANT: Pan, Yang
 ; TITLE OF INVENTION: TANGO-72 AND TANGO-77 NUCLEIC ACID MOLECULES AND POLYPEPTIDES
 ; FILE REFERENCE: MBIO1997-008P1RCP1CNI(M)
 ; CURRENT APPLICATION NUMBER: US/09/862,767A
 ; CURRENT FILING DATE: 2001-05-21
 ; PRIOR APPLICATION NUMBER: US 09/127,856
 ; PRIOR FILING DATE: 1998-08-03
 ; PRIOR APPLICATION NUMBER: US 60/054,646
 ; PRIOR FILING DATE: 1997-08-04
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 9
 ; LENGTH: 458
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-862-767A-9

Query Match 11.2%; Score 370.5; DB 10; Length 458;
 Best Local Similarity 25.5%; Pred. No. 1.2e-21;
 Matches 113; Conservative 80; Mismatches 200; Indels 51; Gaps 11;

QY 119 NLAISOAFESKDSFPKRLIE-----VPYAYQCPYGCASFPGKASGQWEADLH-- 167
 Db 3 NLKRLARSTYNNLKLPTLRLKLVAMEASLTYPHSCAF-----ANMKRQISELHP 54
 QY 168 -----LDESSKRRPGLLARQENHYDQDLDELQLEMEDSKPH--SVQCSPTPG 216
 Db 55 CNKSLIRQEVNDYVTRGGRSSSLAEDNESSYSRGDMTYTEFDYDLCEVVDVCSPKPD 114
 QY 217 PERPCYELFESMGIRLAVMAIVLISVLCNGVILTVFAGGAPLPVKFVVGALAGANTL 276
 Db 115 AENPCDINGYNILRLVIMFISILAITGN--IIVLVITTSQYKLTVPFELMCNLAFADEL 173
 QY 277 TGISGGLASVDALTFQGESEYGARMETGLGCRATGFLAVLGSSEASVLLTLTAAVO--- 332
 Db 174 IGIYLLILASVDIHTKSYHNVAIDMOTGAGCDAAGFTVFASLSVYTLTALTLEHMT 233
 QY 333 --CSVSVSCVRAKGRPSLGSVRAGVLGCLAGLAAALPLASGEYGASPLCLPYAPPE 390
 Db 234 ITHAMOLDC-----KVQLRHAASVMWGMV--FAFAALFEIFGISYSYKVSICLPM--DI 285
 QY 391 GOPPALGFTVALVWMNSFCFLVAGAVIKLYCDLPBGDF--EAVWDCAMVRHVALITADG 449
 Db 286 DSEPLSOLYMSLILVNLVAFVVICGYIHLYVRNNIYSSSDTIARMAALITFD 345
 QY 450 LLYCPVAFISFASMLGLFVTPPEAVSVLLVLPACNLPLLYLLENPHFRDLRLRP 509
 Db 346 LCMAPISFPAISAKLVPLITVSKAKILVLVLFHPINSCANPFLYAIPTKMRDRDFILLS 405
 QY 510 RAGDSGLAYAAAGLEKSSCDST 533
 Db 406 KCG-----CYEMOAOIYRIETST 424

RESULT 14

US-09-862-767A-7
 ; Sequence 7, Application US/09862767A
 ; Patent No. US20020034786A1

; GENERAL INFORMATION:
 ; APPLICANT: Pan, Yang
 ; TITLE OF INVENTION: TANGO-72 AND TANGO-77 NUCLEIC ACID MOLECULES AND POLYPEPTIDES
 ; FILE REFERENCE: MBIO1997-008P1RCP1CNI(M)
 ; CURRENT APPLICATION NUMBER: US/09/862,767A
 ; CURRENT FILING DATE: 2001-05-21
 ; PRIOR APPLICATION NUMBER: US 09/127,856
 ; PRIOR FILING DATE: 1998-08-03
 ; PRIOR APPLICATION NUMBER: US 60/054,646
 ; PRIOR FILING DATE: 1997-08-04
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 7
 ; LENGTH: 359
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-862-767A-7

Query Match 10.6%; Score 350.5; DB 10; Length 359;
 Best Local Similarity 27.5%; Pred. No. 3.4e-20;
 Matches 99; Conservative 68; Mismatches 178; Indels 15; Gaps 6;

QY 175 KRLPGLLARQENHYDQDLDELQLEMEDSKRHPSPVQCSPTPGPCEYLFESMGIRLAV 234
 Db 4 KNPQETLQAFDSHYDYTI-----CGDSE--DMVCTPKSDEFNCPEDIMGYKFLRIV 54
 QY 235 WAIVLLSVLCNGVILTVFAGGAPLPVKFVVGALAGANTLTGISGGLASVDALTGQ 294
 Db 55 WFSYSLALGNVFL--ILTSHKLVNPRFLMCNLAFADECMKMYLLILASVLDLYTHSE 113
 QY 295 FSEYGARMETGLGCRATGFLAVLGSSEASVLLTLTAAVQCSVSVRAVGRKSPISGVRA 354
 Db 114 YYNHAIIDMOTGPGCNAGFTVFASELSVYTLITLERVAITFARMIDKILRHCA 173
 QY 355 GYVGLALAGLAAALPLASGEYGASPLCLPYAPPEGAPALGTVLAVMNSCFELVYA 414
 Db 174 IYGVGWCCFELLALPLVGVISSTAKVSICLPM--DTETPLALAVIVLTLNIVAFYIVC 231
 QY 415 GAVIKLYCDLPBGDFE--AVWDCAMVRHVALIFADGLLYCPVAFISFASMLGPPVPEA 473
 Db 232 CCHVKIYITVRNPNQVNBGDKDTIARMAVLIPTDFICMAPISFYALSAITINKPLITSN 291
 QY 474 VSVLLVLPPLACNLPLLYLLENPHFRDLRLRPRAAGSGPLAYAAAGEL--EKSSCD 531
 Db 292 SKILLVLFYPLNSCANPFLYAIPTKAFOROVFILLSKFGICKRQAOAYRQGRVPPKNSD 351

RESULT 15

US-09-862-767A-5
 ; Sequence 5, Application US/09862767A
 ; Patent No. US20020034786A1
 ; GENERAL INFORMATION:

; APPLICANT: Pan, Yang
 ; TITLE OF INVENTION: TANGO-72 AND TANGO-77 NUCLEIC ACID MOLECULES AND POLYPEPTIDES
 ; FILE REFERENCE: MBIO1997-008P1RCP1CNI(M)
 ; CURRENT APPLICATION NUMBER: US/09/862,767A
 ; CURRENT FILING DATE: 2001-05-21
 ; PRIOR APPLICATION NUMBER: US 09/127,856
 ; PRIOR FILING DATE: 1998-08-03
 ; PRIOR APPLICATION NUMBER: US 60/054,646
 ; PRIOR FILING DATE: 1997-08-04
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 5
 ; LENGTH: 458
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-862-767A-5

Query Match 9.6%; Score 319; DB 10; Length 458;
 Best Local Similarity 24.9%; Pred. No. 1.4e-17;
 Matches 103; Conservative 77; Mismatches 213; Indels 20; Gaps 8;

Sat Nov 9 13:04:29 2002

us-09-851-595-5.ra1

Page 1

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2002, 19:33:15 : Search time 12.9467 seconds
(Without alignments)
1438.571 Million cell updates/sec

Title: US-09-851-595-5

Perfect score: 3307
Sequence: 1 NTHTYRESWACRYRSGIPG.....GGLSGGGGFGPQSLAFASHV 633

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:.*
5: /cgn2_6/ptodata/1/1aa/PCPNUS.COMB.pep:.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	958	29.0	644	2	US-08-866-757-2	Sequence 2, Appl
2	958	29.0	644	4	US-09-153-593-2	Sequence 2, Appl
3	419.5	12.7	764	4	US-07-741-453A-60	Sequence 60, Appl
4	419	12.7	696	4	US-07-757-342D-6	Sequence 4, Appl
5	417	12.6	692	4	US-07-757-342D-6	Sequence 6, Appl
6	414.5	12.5	695	3	US-08-487-886-2	Sequence 2, Appl
7	414.5	12.5	695	3	US-08-482-855-2	GENERAL INFORMA
8	414.5	12.5	695	4	US-08-474-986-2	Sequence 54, Appl
9	414.5	12.3	764	4	US-07-741-453A-54	Sequence 3, Appl
10	407.5	12.3	764	4	US-07-757-342D-3	Sequence 61, Appl
11	407.5	12.3	764	4	US-07-741-453A-61	Sequence 5, Appl
12	406.5	12.3	764	4	US-07-757-342D-5	Sequence 59, Appl
13	403.5	12.2	674	4	US-07-741-453A-59	Sequence 10, Appl
14	381	11.5	674	4	US-07-757-342D-10	Sequence 2, Appl
15	381	11.5	699	4	US-07-757-342D-2	Sequence 5, Appl
16	366.5	11.1	792	4	US-07-741-453A-56	Sequence 8, Appl
17	356	10.8	795	4	US-07-741-453A-55	Sequence 55, Appl
18	354	10.7	611	4	US-07-757-342D-8	Sequence 7, Appl
19	354	10.7	636	4	US-07-757-342D-7	Sequence 29, Appl
20	343.5	10.4	764	4	US-07-741-453A-29	Sequence 33, Appl
21	318.5	9.6	420	4	US-08-795-876-33	Sequence 38, Appl
22	318.5	9.6	423	4	US-08-795-876-38	Sequence 2, Appl
23	318.5	9.6	436	4	US-08-795-876-2	Sequence 34, Appl
24	313	9.5	336	1	US-08-118-270-54	Sequence 54, Appl
25	313	9.5	336	5	PCR-US93-08528-54	Sequence 55, Appl
26	289.5	8.8	327	1	US-08-118-270-55	Sequence 55, Appl
27	289.5	8.8	327	5	PCT-US93-08528-55	Sequence 55, Appl

28	250.5	7.6	332	1	US-08-118-270-53	Sequence 53, Appl
29	250.5	7.6	332	5	PCT-US93-08528-53	Sequence 53, Appl
30	140	4.2	673	4	US-09-063-950-2	Sequence 2, Appl
31	135.5	4.1	370	4	US-08-900-230-2	Sequence 2, Appl
32	135.5	4.1	370	4	US-08-900-230-2	Sequence 2, Appl
33	134.5	4.1	369	4	US-09-172-353-6	Sequence 6, Appl
34	134.5	4.1	370	4	US-09-199-737-2	Sequence 8, Appl
35	134.5	4.1	372	2	US-08-626-685A-8	Sequence 2, Appl
36	134.5	4.1	372	4	US-08-993-088A-2	Sequence 2, Appl
37	134.5	4.1	372	4	US-08-993-424B-2	Sequence 2, Appl
38	134.5	4.1	372	4	US-08-665-034A-4	Sequence 4, Appl
39	134.5	4.1	371	4	US-08-993-088A-9	Sequence 9, Appl
40	134	4.1	371	4	US-08-993-424B-9	Sequence 9, Appl
41	134	4.1	350	2	US-08-466-103A-12	Sequence 12, Appl
42	131	4.0	353	2	US-08-466-103A-12	Sequence 6, Appl
43	131	4.0	361	3	US-09-034-985-2	Sequence 2, Appl
44	131	4.0	361	3	US-09-034-985-2	Sequence 2, Appl
45	130.5	3.9	372	4	US-08-993-088A-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-08-866-757-2
Sequence 2, Application US/08866757
Patent No. 5858716
GENERAL INFORMATION:
APPLICANT: ELISHOURANGY, NABIL A
APPLICANT: LI, XIAOTONG
APPLICANT: BERGMA, DEK J
TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAAT1)
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/866,757
FILING DATE: 30-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70055
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 644 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-866-757-2
Query Match 29.0% Score 958; DB 2; Length 644;
Best Local Similarity 36.6% Pred. No. 7.8e-76;
Matches 211; Conservative 85; Mismatches 191; Indels 60; Gaps 9;

```

QY 54 KLEADTFQSLSIALDLDSWMAIRSHPEARSTLHSLVKLDTJNQTLTTLPLAGLGIMH 113
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 72 QIKETFGGLSLRLDLSRNLHIEHISRAFAATLGPITNLPVSEFNEILTSFTEGLINGLNO 131
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 114 LKLGKLNALSAFSDSPKRLILEVPAYOCCPYGMCASFEEKASGQWEADLHLD--- 170
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 132 LKLGKLNALSAFSDSPKRLILEVPAYOCCPYGMCASFEEKASGQWEADLHLD--- 170
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 171 -EESKRRLGLLARAENHHYDDELQLEMDSKRPHSVOCSPPTGPFKCEYLFEESWG 229
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 188 AOEKGTADAANVTSTLENEHSOI-----IHCPTSGAFKCEYLLGSMW 233
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 230 IRLAVMAIVLTLVLCNGVLTLTFVAGGAPLPPVKFVGAAGANTLTGISCGLIASVDA 289
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 234 IRLTWMFIFLVALFNLVLTITTFASCT--LPSSKLEIGLISVNLPMGIYTGILTFDA 292
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 290 LTFGPFSEYGARMTGICGRATGFLAVLGEASVLLTLAAVOCVSVCRAAGKPSL 349
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 293 VSMGRFAEFIMWETGSGCKVGTGLAVFSESSEAFILMLATVERSLAKIMKNGSNHL 352
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 350 GSVRAGVLCGLALAGLAAALPLASVGEYASPLCLPYAPPEGOPALGFTVALVMMNSFC 409
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 353 KOPRVAALAFIATVTCGCPFLPHRGEYSASPLCLP--PTGETPSLGFTVTLVLSLA 410
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 410 FLVAVAGATIKYCDLPRGDFEAVMDCAMVRHVAWLIFADGILYCPVAFISFASMLGLFVY 469
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 411 FLMAVITKLYCNLEKEDLSSENSOSSMIKHVAWLIFTCIFCCVAFESFAPLITAI 470
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 470 TPEAVKSVLLVPLPACINPLVLTLEPNFRDD---LRRLPRAGDSGLAYAAAGEL 525
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 471 SPEIMKSVTLIFPLPACINPLVLTLEPNFRDD---LRRLPRAGDSGLAYAAAGEL 525
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 526 EK-----SSCDSTQALVAFSVD---LLEASERAGRPGLTETVGFPSVT 566
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 531 EODEFYDDGMSHLOGNLTVCDCESPFLTRKPVSKHLI-----KSHSCPALA 578
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 567 LISCOOP 573
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 579 VASCOOP 585
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 2
US-09-153-593-2
; Sequence 2, Application US/09153593A
; Patent No. 6174994
; GENERAL INFORMATION:
; APPLICANT: ELISHOURBAGY, NABIL A
; APPLICANT: LI, XIAOTONG
; TITLE OF INVENTION: NOVEL 7PM RECEPTOR (H2CAAT1)
; FILE REFERENCE: GH-70055-1
; CURRENT APPLICATION NUMBER: US/09/153, 593A
; EARLIER FILING DATE: 1998-09-15
; EARLIER APPLICATION NUMBER: 08/866, 757
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 644
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-153-593-2

```

```

Query Match 29.0%; Score 958; DB 4; Length 644;
Best Local Similarity 38.6%; Pred. No. 78e-76;
Matches 211; Conservative 85; Mismatches 191; Indels 60; Gaps 9;
QY 54 KLEADTFQSLSIALDLDSWMAIRSHPEARSTLHSLVKLDTJNQTLTTLPLAGLGIMH 113
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 72 QIKETFGGLSLRLDLSRNLHIEHISRAFAATLGPITNLPVSEFNEILTSFTEGLINGLNO 131
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 114 LKLGKLNALSAFSDSPKRLILEVPAYOCCPYGMCASFEEKASGQWEADLHLD--- 170
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db 132 LKLGKLNALSAFSDSPKRLILEVPAYOCCPYGMCASFEEKASGQWEADLHLD--- 187
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 171 -EESKRRLGLLARAENHHYDDELQLEMDSKRPHSVOCSPPTGPFKCEYLFEESWG 229
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 188 AOEKGTADAANVTSTLENEHSOI-----IHCPTSGAFKCEYLLGSMW 233
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 230 IRLAVMAIVLTLVLCNGVLTLTFVAGGAPLPPVKFVGAAGANTLTGISCGLIASVDA 289
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 234 IRLTWMFIFLVALFNLVLTITTFASCT--LPSSKLEIGLISVNLPMGIYTGILTFDA 292
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 290 LTFGPFSEYGARMTGICGRATGFLAVLGEASVLLTLAAVOCVSVCRAAGKPSL 349
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 293 VSMGRFAEFIMWETGSGCKVGTGLAVFSESSEAFILMLATVERSLAKIMKNGSNHL 352
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 350 GSVRAGVLCGLALAGLAAALPLASVGEYASPLCLPYAPPEGOPALGFTVALVMMNSFC 409
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 353 KOPRVAALAFIATVTCGCPFLPHRGEYSASPLCLP--PTGETPSLGFTVTLVLSLA 410
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 410 FLVAVAGATIKYCDLPRGDFEAVMDCAMVRHVAWLIFADGILYCPVAFISFASMLGLFVY 469
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 411 FLMAVITKLYCNLEKEDLSSENSOSSMIKHVAWLIFTCIFCCVAFESFAPLITAI 470
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 470 TPEAVKSVLLVPLPACINPLVLTLEPNFRDD---LRRLPRAGDSGLAYAAAGEL 525
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 471 SPEIMKSVTLIFPLPACINPLVLTLEPNFRDD---LRRLPRAGDSGLAYAAAGEL 525
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 526 EK-----SSCDSTQALVAFSVD---LLEASERAGRPGLTETVGFPSVT 566
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 531 EODEFYDDGMSHLOGNLTVCDCESPFLTRKPVSKHLI-----KSHSCPALA 578
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 567 LISCOOP 573
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 579 VASCOOP 585
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 3
US-07-741-453A-60
; Sequence 60, Application US/07741453A
; Patent No. 6228597
; GENERAL INFORMATION:
; APPLICANT: PARMENTIER, MARC
; APPLICANT: LIBERT, FREDERIC
; APPLICANT: DUMONT, JACQUES
; APPLICANT: VASSART, GILBERT
; TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR
; TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1615 L STREET, N.W.
; CITY: WASHINGTON, D.C.
; COUNTRY: U.S.A.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 1991/015
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16773
; REFERENCE/DOCKET NUMBER: 91913/1107/US/ST
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:

```

Gaps 12

211

THE JST

HY 189
1

LC 244

ET 304

AC 364

493

AV 553

YP 613

```

APPLICANT: IGARASHI, Masao
            MINEGISHI, Takashi
            NAKAMURA, Kazuo
TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
            CUSHMAN
            STREET: 130 Water Street
            CITY: Boston
            STATE: Massachusetts
            COUNTRY: US
            ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,342D
FILING DATE: 10-Sep-1991
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BUCKLEY, Linda M.
REGISTRATION NUMBER: 31003
REFERENCE/DOCKET NUMBER: 41226
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)523-3400
TELEFAX: (617)523-6440
TELEX: 200291 STRE UR
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 692 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-07-757-342D-6
Query Match      12.6%  Score 417; DB 4; Length 692;
Best Local Similarity 25.0%  Pred. No. 3.6e-28;
Matches 167; Conservative 86; Mismatches 251; Indels 164; Gaps 23;
QY 16 SGIPSTHASVRSQGL-----SLPAHPSLAALAASTTASGKLEXTDFQSLSL 66
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 92 SNLPKRLHRIEKANLLYINPEAFQNP-----SLRYLLISNT---GIKHLPAVHKIQSL 144
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 67 Q-----ALDLSMNAIRSIHPEAF-----S 85
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 145 QKVLDDIODNINIHIVANNSFMGISESVIIMLSKNGIEEIHNCADFNGTQDELNLSDNN 204
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 86 TLHSL-----VKLDLTNDQLTTLPLAGLGMLKLGKGNLALSOAKSDFPKL 134
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 205 NLEELPNDVFOGASGPVILDISRTKRVHSLPHNGLENLKKLRARSTYRLKKLPNLDKFEVTL 264
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 135 RILEVPAVYQCCPRGKASGQWAEEDLH-----LD 169
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 265 MEASLTPSHCCAF-----ANLKRQISELHPICKNSILRODIDMTQIGDQVSLSD 316
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 170 DEESSKRLGLLARAQENHYDQDELQLEMEDESKRHPVSQCSPTGPKREYLFESNG 229
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 317 DEPS-----YKGSDDMMNNEFDYDCN---EYVD-----VTCSPKPDAFNCEDEIMGYNI 363
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 230 IRLAVALIYLSVLCNGVLVLLTFVAGAPAPLPVKFYVGAIGAGANTLTGISCGLIASYDA 289
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 364 LRVLIWFISLITATGNTTVLV-VLTTSQYKLVPRFLMNCNLAFAFDLCIGIYLLIASYDI 422
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 290 LTFGQSEYGARWETGCGATGFLAVLGSEASYLLTLAAVQ-----CSVSVCYRAY 343
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 423 HTKSQYHNVAIDMOTGACCAAGFTVFASLSYTLTATLEWHITTHAMQLECC----- 478
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 344 GSPSLGAVRAGVLCIALAGLAALPLASVGEYGASPLDLPVAPPGCAPALGFYVALV 403
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Db 479 -KVQLRHAASVMVLG-WTFAPAAALFPIFGISYKKSICLPM--DIISPLSQLYVMALL 534
QY 404 MNNSFCFLVYAGATIKLYCULPRGDF-EAYVDCAMVRHVAMLIRADGLLYCPVAFISFAS 462
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 535 VLNVLAFFVVGCGTYHLYLVNRPNTIVSSSDTIKAKRMATLIFTDLCMPISFFALSA 594
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 463 MGLGFPVTPPAVKSVLVLPPLPACLPILYLLFNPHRDRLRLRRAGDSGLPATAAA 522
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 595 SLKVPITTVSKAILLVLFYPINSCANPFLAITTKNRRDFTL----- 639
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 523 GELEKSSCDSTQALVAFSDVDLLIEASAGRPGLTTPGPPSVLLISQOQCAPRLGSH 582
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 640 --LSKFCGYEMOAOIYRT-----ETSSA-----THNF-HARKSICS--SAPRVVNSY 681
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 583 CVPBEGNH 590
    | | |
Db 682 VLVP-LNH 688
    | | |
RESULT 6
US-08-487-886-2
; Sequence 2, Application US/08487886
; Patent No. 5744448
GENERAL INFORMATION:
APPLICANT: Kelton, Christie Ann
APPLICANT: Schweickhardt, Rene Lynn
APPLICANT: Cheng, Shirley Vui Yen
APPLICANT: Nugent, No. 5744448een Patrice
TITLE OF INVENTION: Human Follicle Stimulating
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: Stephan P. Williams,
ADDRESS: Ares-serono, Inc.
STREET: Exchange Place, 37th floor
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
COMPUTER: IBM PS/2, model 55 SX
OPERATING SYSTEM: MS-DOS version 4.0
SOFTWARE: VAX/VMS Massll via Kermit to IBM MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,886
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/670,085
FILING DATE: 15-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Stephan P.
REGISTRATION NUMBER: 28546
REFERENCE/DOCKET NUMBER: US/252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 723-1300
TELEFAX: (617) 723-8923
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 695
TYPE: Amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: signal sequence
LOCATION: -17 to -1
IDENTIFICATION METHOD: hydrophobic
NAME/KEY: putative amino-terminal extracellular domain
LOCATION: 1 to 349
IDENTIFICATION METHOD: similarity with other
IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular

```

IDENTIFICATION METHOD: domains, hydrophilic
FEATURE:
NAME/KEY: transmembrane domain
LOCATION: 350 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains
FEATURE:
NAME/KEY: putative transmembrane region I
LOCATION: 350 to 370
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region II
LOCATION: 382 to 404
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region III
LOCATION: 427 to 448
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region IV
LOCATION: 469 to 491
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region V
LOCATION: 512 to 533
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VI
LOCATION: 557 to 580
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VII
LOCATION: 592 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative carboxy-terminal intracellular
LOCATION: 614 to 678
US-08-487-886-2

Query Match 12.5% Score 414.5; DB 1; Length 695;
Best Local Similarity 25.0%; Pred. No. 6e-28;
Matches 131; Conservative 94; Mismatches 232; Indels 67; Gaps 13;

55 LENDPFSOLS-SLQALDLSWNAIRSIHPEAF-----STLHSL----- 90
160 IENNSVGLSFESVILMLNKNGIOETINCAFNQDLELNSDNNNEELPNDVPHGASG 219
91 -VKLDLTNDQULTPLAGLGLMLHLKGNLAISQAFSKSPFLRIIEVPAYOCPPYG 149
220 PYLIDISRTIRHISLPSGLENLKLRLARSTYNNLKLPTLEKLVAMEASTYPSHCAF- 278
150 MCASFPKASQWAEADH-----LDDDESSRPLGLLAROENHYDODDDELQOL 198
279 -----AMRRROISELHPICNKSILRQEVDMYTQTRGQSSSLAEDNESSYSRQFDMYTT 331
199 EMDSPKHP--SVQCSPTPQPFKPCCEYLFEWSGIRLVAWVALVLSVLCNGVLTLVFAGG 256
332 EFYDLCNEVVDVTGSKPAPAFNCPEDIMGVNIILRVLWFSILAIQGN-ILIVLILTS 390

257 PAPDPKFEVGAIGANTLTIGSCGLASVDALFFGQFSEYGARFETGLCRATGFLAV 316
391 QYKLVPRFLMGNLAFADICIGIYLLIASVDIHFTSQXHNVAIDQTSAGCDAAEFVY 450
317 LGSEASVLLTLTAAYQ-----CSVSVSCVRAVGRKPSISGVSAGVGLGCLALAGLAAAP 370
451 FASELSVYTLTAITLERWHTITHAMQDC-----KQLRHNASVWVGWT-FAFAALFP 504
371 LASVEYGASPLCLPYAPEGPPALAGFTVALVMNSFCFLVAGAYIKLYCDLPBGDF- 429
505 IFGISYMKVSTCLPM--DIDSPLSQLYMSLLVNLVLAFFVIGCGYIHLYLVARNPIY 562
430 EAVWDCAMRVHVAMLIFADGLYCPVAPLSPASMLGLFPVTPPAVKSVLVYLPLPACLN 489
563 SSSSTRLAKRRAMLIETDFLCMAPISFAISASLKVPLITVSKAKILLVLFHINSKAN 622
490 PLLYLLENFPRDDLRRLPRAGDSGLAVYAAAGCEKSDCST 533
623 PLVIAIFTKNFRDFILLKCG-----CYEMQAQILYRTETST 661

RESULT 7
US-08-482-855-2
Sequence 2, Application US/08482855
Patent No. 6121016
GENERAL INFORMATION:
APPLICANT: Kelton, Christie Ann
APPLICANT: Schweickhardt, Rene Lynn
APPLICANT: Cheng, Shirley Vui Yen
APPLICANT: Nugent, No. 6121016en Patrice
TITLE OF INVENTION: Human Follicle Stimulating
TITLE OF INVENTION: Hormone Receptor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Stephan P. Williams,
ADDRESSEE: Ares-Serono, Inc.,
STREET: Exchange Place, 37th floor
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
COMPUTER: IBM PS/2, model 55 SX
OPERATING SYSTEM: MS-DOS version 4.0
SOFTWARE: VAX/VMS Massil via Kermit to IBM MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482, 855
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/670, 085
FILING DATE: 15-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Stephan P.
REGISTRATION NUMBER: 28546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 723-1300
TELEFAX: (617) 723-8923
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 695
TYPE: Amino acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: signal sequence
LOCATION: -17 to -1
IDENTIFICATION METHOD: hydrophobic
FEATURE:
NAME/KEY: putative amino-terminal extracellular domain

LOCATION: 1 to 349 similarity with other
IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular
IDENTIFICATION METHOD: domains, hydrophilic
FEATURE:
NAME/KEY: transmembrane domain
LOCATION: 350 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains
FEATURE:
NAME/KEY: putative transmembrane region I
LOCATION: 350 to 370
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region II
LOCATION: 382 to 404
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region III
LOCATION: 427 to 448
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region IV
LOCATION: 469 to 491
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region V
LOCATION: 512 to 533
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VI
LOCATION: 557 to 580
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VII
LOCATION: 592 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative carboxy-terminal intracellular
NAME/KEY: domain
LOCATION: 614 to 678
US-08-482-855-2

Query Match 12.5%; Score 414.5; DB 3; Length 695;
Best local Similarity 25.0%; Pred. No. 6e-28;
Matches 131; Conservative 94; Mismatches 232; Indels 67; Gaps 13;

55 LEXDFTSOLS-SLOALDSWMAIRSIHPEAF-----STLHSL----- 90
DB 160 IERNSEFGLSFESEVILMLNKNGIOEIHNCFAENGTOLELNLSDNNLLELPNDVFGASG 219
QY 91 -VKLDLDLNDQLTPLPLAGLGLMLHLKGNLALSAFSDSPKRLITLPPAYVQCCPYG 149
DB 220 PVLLDISRTIHSHSPSYGLKLEKRLARSTYLNKLPLEKLVALMEASITYPSSHCAF- 278
QY 150 MCASFASGCOMWEADLH-----LDDESSKRPGLGLARQENHYDDLDLEQL 198
DB 279 -----AMRRQISELHPICNKSILROEVDMYOTRGORSSLAEDNESSYSKGFMTYT 331

QY 199 EMEDESKPH--SVQSPPTPGPKPCCEYLFEWSGIRLAWAIVLVLSCNGVLTLPVAGG 256
DB 332 EFDYDLCNEVVDYTCSPKPAFNPCEIDIMGNILRVLWIFSLIATIGN-ITVYILTS 390
QY 257 PAPLPVKRFVVGAIAGANTLTGISCGLLASVDALTFQOSEYGARWETGLCCRATGFLAY 316
DB 391 QYKLVPRFLMCLNLAADLCIGIYLLILIASVDITHKSOYHNVAIDWQGACGDAAFFTV 450
QY 317 LGSEASVLLTLAAVQ-----CSVSYCYRAVKSLSLGSVRAGVLCGLAAGLAALP 370
DB 451 FASELSVTLTALTTLERWHTTHAMQDLC---KVQLRHAASVVMGM-FAFAALLP 504
QY 371 LASVGEYASPLCLPAPPEGPAALGFTVALVMNSCFPLVAGAYKLYCDDLPBGDF- 429
DB 505 IFCISSYMYVSLCPM--DIDSPLSQLYVMSLVLVNLAFLVYICGYIHITLVANPNIV 562
QY 430 EAVWDCAMVRHVAWLIFADGLLYCPVAFLSPASMLGLEPVTPEAVKSVLVLPPLACLN 489
DB 563 SSSSPRIAKRMAMLIETDPLCMAPISFPAISALKPLITVSKAKILLVLFHPINSCAN 622
QY 490 PLLYLFFNHFRRDLRLRRADSGPLAAYAGLEKSSCDST 533
DB 623 PFLYALFTKFRFRPFLILSKG---CYEMQAQIVRTTST 661

RESULT 8
US-08-474-986-2
GENERAL INFORMATION:
APPLICANT: Kellon, Christie Ann
Schweickhardt, Rene Lynn
Cheng, Shirley Vui Yen
Nugent, No. 637271leen Patrice
TITLE OF INVENTION: Human follicle stimulating
Hormone Receptor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSER: Stephan P. Williams,
Ares-Serono, Inc.
STREET: Exchange Place, 37th floor
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
COMPUTER: IBM PS/2, model 55 SX
OPERATING SYSTEM: MS-DOS version 4.0
SOFTWARE: VAX/VMS Mass11 via Kermit to IBM MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474, 986
FILING DATE: 07-Jun-1995
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/670, 085
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Stephan P.
REGISTRATION NUMBER: 28546
REFERENCE/DOCKET NUMBER: US/252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 723-1300
TELEFAX: (617) 723-8923
LOCATION: 614 to 678
SEQUENCE DESCRIPTION: SEQ ID NO: 2;
US-08-474-986-2

Query Match 12.5%; Score 414.5; DB 4; Length 695;
Best local Similarity 25.0%; Pred. No. 6e-28;
Matches 131; Conservative 94; Mismatches 232; Indels 67; Gaps 13;

55 LEXDFTSOLS-SLOALDSWMAIRSIHPEAF-----STLHSL----- 90
DB 160 IERNSEFGLSFESEVILMLNKNGIOEIHNCFAENGTOLELNLSDNNLLELPNDVFGASG 219

RESULT 10
 US-07-757-342D-3
 Sequence 3, Application us/07757342D
 Patent No. 6218509
 GENERAL INFORMATION:
 APPLICANT: IGARASHI, Masao
 MINEGISHI, Takashi
 NAKAMURA, Kazuto
 TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: DAVID G. COMLIN, DIKE, BRONSTEIN, ROBERTS &
 CUSHMAN
 STREET: 130 Water Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: US
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,342D
FILING DATE: 10-Sep-1991
CLASSIFICATION: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BUCKLEY, Linda M.
REGISTRATION NUMBER: 31003
REFERENCE/DOCKET NUMBER: 41226
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)523-3400
TELEFAX: (617)523-6440
TELEX: 200291 STRE UR
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 700 amino acids
TYPE: amino acid
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-07-757-342D-3

```

```

Query Match
Best Local Similarity 12.3%; Score 407.5; DB 4; Length 700;
Matches 135; Conservative 99; Mismatches 254; Indels 75; Gaps 13;

```

```

QY 47 SNTTSGKLEXPDSFSSQSS-----LQALDLSWNA-INSIHPEAFSTLSVKLDLTDN 98
DB 177 NNESTYTKLYGFEVQSHANGNTLLISLEKENYLEKMHSGAFQAGATGSILDISST 236
QY 99 QLTTPPLAGLGLMLHKLKGNLALSOAFSKDSFPKRLIEVPYAVOCCPYGMCASFRRAS 158
DB 237 KLOALPSHGLSIOFLILASSYSLTKLPSEKKEFTSLVATLVTPSHCCAFNRLLPK----- 291
QY 159 GQWEAEDHLDEESSKRPGLGLARQAEH--YDQDDELQLEMED-----SKRPSYOC 211
DB 292 ---KEONFSFIEFMSQCESTYKAKNETLYSAIFEENELSGMDYDYGCSF-KTIQC 347
QY 212 SPTGPRPCPYLEESGIRLAVALVLSVLCNGVLVTFAGGAPAPLPYKPVYGAIA 271
DB 348 APEPDARNPCEIDINGYAFRLVILWILNITLIFGN-LTVFLVLLSRKLYTPRFLMCLMS 406
QY 272 GANTLTGISCGLASVALTFGQSEFGARWETGLCRATGFLAVLGSASVLLTLTAAY 331
DB 407 FADFCMGLYLLLASVDSQTKGQYNNHAIQWQSGCGAAGFTVFASSELVYTLTVTL 466
QY 332 QCSVSVCAVAGKSPSLGVRAGVGLCLAGLAALPLASVGEYGASPLCLPYARPEG 391
DB 467 ERNHTTYVAVQDQKRLRLRAIPIMLGGWLFSTLIATMPLVGISNMYKVSICLPM--DVE 524
QY 392 QPAAIGFTVALVMNSFCFLVAVGAYIKLYCDLPRGDFEAV-WDCANVRHVAWLIFADGL 450
DB 525 STLSQYIILSILNLVAVVICACIIRITFAVONPELTAPNNDTKRIAKKMAILIFDTF 584
QY 451 LVCYPAFLSFASMLGLFPYTPPEAVKSVLLVLPACILNPLVLLNPHRDLRLRLR 510
DB 585 CMAPISFEFAISAFKPLVTLVNSKILLVFPVNSCANPFLAITTKAFORFELLILSR 644
QY 511 AGDSGPLYAAAGLEKSCDSTOALVAFSDVDLIEASEAGRPQLEY-----GFP- 563
DB 645 FG-----CCKRRA-----ELVRRKEFSAYTSNCKNGPFG 673
QY 564 -----SVTLSCQGPAGAPR 577
DB 674 ASKPSQATLKLTIVHCOQPIPR 696

```

```

RESULT 11
US-07-741-453A-61
; Sequence 61 Application US/0741453A
; Patent No. 6228597
; GENERAL INFORMATION:
; APPLICANT: PARMENTIER, MARC

```

```

APPLICANT: LIBERT, FREDERIC
APPLICANT: DUMONT, JACQUES
APPLICANT: VASSART, GILBERT
TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR
TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DAREY & CUSHMAN
STREET: 1615 L STREET, N.W.
CITY: WASHINGTON, D.C.
COUNTRY: U.S.A.
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/741,453A
FILING DATE: 19911015
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KORULIS, PAUL N.
REGISTRATION NUMBER: 16773
REFERENCE/DOCKET NUMBER: 91913/1107/US/ST
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 764 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
MOLECULE TYPE: protein
US-07-741-453A-61

```

```

Query Match
Best Local Similarity 12.3%; Score 407.5; DB 4; Length 764;
Matches 145; Conservative 99; Mismatches 261; Indels 105; Gaps 13;

```

```

QY 19 PGSTHASYERSQGLSLPAHPRASIALAASNTTASGKLEXPDSQ-----SSIQALDL 71
DB 162 PYMTSIPVNAFQGL-----CNETLTCLKYNGFTSVQCYAFNGTKLDIAVYL 207
QY 72 SWNA-INSIHPEAFSTLSLVK-LDLTDNQLTTPPLAGLGLMLHKLKGNLALSOAFSKD 129
DB 208 NKKKYLTVLIDKDAFGGYSGPSLIDVQTSVTLAPSKGLHKLKELIRKNTWTLLKPLSL 267
QY 130 SFPKRLIEVPYAVOCCPY-----GMCASF----- 154
DB 268 SFLHTRADLSTYSHCCAFNRKKNKIGLIESLNCNESSMOSLRQKRSVNALNSPLHQEYE 327
QY 155 -----FKASGV-----EADDLHLDEESSKRPGLGLARQAEHNYDQ 191
DB 328 ENLGDSTVGYKSKSPQDTHNNAHYVFFPEQDEDLIGGQELAKNQDETLQAFDSHYD 387
QY 192 DDELQLEMEDSKRPHSVOCSPTPRGFKCEYLFESWGIRLAVALVLSVLCNGVLVLT 251
DB 388 TL-----CGDSE---DWCITKSPDEFNCEIDINGYKFLRIYVWVFSLLALGNFVLL- 437
QY 252 VFAGGPAPLPYKPVYGAIAAGANTLTGISCGLASVDALTFGQSEFGARWETGLCRAT 311
DB 438 ILLTSHYKLNVPRLMCLNDAFDFCMGMYLLILASVDLTYTSEYNNHAIQWQSGCWTA 497
QY 312 GLFVAVGSEASVLLTLTAAVOCSSVSCVRAVYKSPSGSVAGVGLCLAGLAALPL 371
DB 498 GFTVFASELSVYTLVTLTTERWYAITFAMRLDKRIRLRHAAAIWVGWCCFLALPL 557
QY 372 ASVEYGASPLCLPYAPPEGOPALGFTVALVMNSFCFLVAVGAYIKLYCDLPRGDFE- 430

```

Db 558 VGISSYAKVSIQPM--DTEPLALAYIMSVLNIYAFVIVCCYKITYIVRNPOYN 615

QY 431 AVMOCAMRVHVAWMLFADGLYCVAFLSFASMLGFPYTPBAVKSULLVPLPACLNP 490

Db 616 GDKDTKIKRAKRAVILFTDFICMAPISFVLSALINKPLITVSNKILLYLFVPLNSCAMP 675

QY 491 LLYLLEPHFRDRLRLPRAGDSGLPAYAAG-----ELEKSSCDSTQALVAF 539

Db 676 FLYAIFTKAFQRDVFILSKFGICKROAQAYRGQVPRKNSDIDIOVKYTHDRGGLHNM 735

QY 540 SDVDLLEAS 549

Db 736 EDVYELLENS 745

RESULT 12

US-07-757-342D-5

Sequence 5, Application US/07757342D

Patent No. 6218509

GENERAL INFORMATION:

APPLICANT: IGARASHI, Masao

NAKAMURA, Kazuo

MINEGISHI, Takashi

TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: DAVID G. CUSHMAN, DIKE, BRONSTEIN, ROBERTS & CUSHMAN

STREET: 130 Water Street

CITY: Boston

STATE: Massachusetts

COUNTRY: US

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/757,342D

FILING DATE: 10-Sep-1991

CLASSIFICATION: <unknown>

ATTORNEY/AGENT INFORMATION:

NAME: BUCKLEY, Linda M.

REGISTRATION NUMBER: 31003

REFERENCE/DOCKET NUMBER: 41226

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)523-3400

TELEFAX: (617)523-6440

TELEX: 200291 STRE UR

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 764 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-07-757-342D-5

Query Match 12.3%; Score 406.5; DB 4; Length 764;

Best Local Similarity 23.8%; Pred. No. 3.5e-27;

Matches 145; Conservative 98; Mismatches 262; Indels 105; Gaps 13;

QY 19 PGSTHASVRSQGLSLPAHPASLAALASNTTASGKLENDTFQSOL-----SLQALDL 71

Db 162 PYMTSLPVNAFQGL-----CMETLTIKLYNNGFTSVQGYAFNGTKLDAVYL 207

QY 72 SWNA-IRSHIPEARSTLHSLVK-LDLTDNOLITPLAGLIGIMHLKLGKGNALISQASKD 129

Db 208 NKNXYLVIAIKDAFGVSGPSLDVDSQTSVTALPSKGLHLKELIARNWTWTLKKLPISL 267

QY 130 SFPRRLLEVPYAVQCCPY-----GMCASF----- 154

Db 268 SFHLHLRADLSYPSHCCAFKRNQKIRGLIESLMCNSSMOSLRQKSVNALSPLHQYE 327

QY 155 -----FKASGQW-----EADHLHDEESSKRPLGLLARQENHHYQ 191

Db 328 ENIGDSIVGKREKSKFODHNNNAHYVFFEEQDEILIGQELKNFOEFTLQAFDSHYD 387

QY 192 DLDELQLEMEDSKPHPSVOCSPPIPGPKCEYLFESMGRLAVWAVILSVLNGVLLT 251

Db 388 TI-----CGDSE---DMWCTPKPSDEPNPCEDIMGKFLNIVWVYSLALLGNVFL- 437

QY 252 VFAGCAPLPPYKRVYGAIGANTLTISCGLLASVDALTFGQFSEYGARWETGLGRAT 311

Db 438 ILTSHYKLVNPRFLMGNLAFADRCMGNVLLILASVDLYTHSEYNNHAIIDMOTGPCNTA 497

QY 312 GFLAVLGSEASVLLTLTAAVQSVSVCRAVYKSPSGSVRAGVGLCLAGLAALPL 371

Db 498 GFTVFASSELVTLVITTELRWYAITFAMRLDRKMLRHACALVMGWCCFLALPL 557

QY 372 ASVGEYASPLCLPYAPPEQPALGFTVALVMNNSCFVYVAGAYIKLYCDLPRGDFE- 430

Db 558 VGISSYAKVSIQPM--DTEPLALAYIVLNIYAFVIVCCYKITYIVRNPOYN 615

QY 431 AVMOCAMRVHVAWMLFADGLYCVAFLSFASMLGFPYTPBAVKSULLVPLPACLNP 490

Db 616 GDKDTKIKRAKRAVILFTDFICMAPISFVLSALINKPLITVSNKILLYLFVPLNSCAMP 675

QY 491 LLYLLEPHFRDRLRLPRAGDSGLPAYAAG-----ELEKSSCDSTQALVAF 539

Db 676 FLYAIFTKAFQRDVFILSKFGICKROAQAYRGQVPRKNSDIDIOVKYTHDRGGLHNM 735

QY 540 SDVDLLEAS 549

Db 736 EDVYELLENS 745

RESULT 13

US-07-741-453A-59

Sequence 59, Application US/07741453A

Patent No. 6228597

GENERAL INFORMATION:

APPLICANT: PARMENTIER, MARC

LIBERT, FREDERIC

APPLICANT: DUMONT, JACQUES

APPLICANT: VASSART, GILBERT

TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR

TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS

NUMBER OF SEQUENCES: 62

CORRESPONDENCE ADDRESS:

ADDRESSEE: CUSHMAN DARBY & CUSHMAN

STREET: 1615 L STREET, N.W.

CITY: WASHINGTON, D.C.

COUNTRY: U.S.A.

ZIP: 20036

COMPUTER READABLE FORM:

MEDIUM TYPE: Disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/741,453A

FILING DATE: 19911015

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: KOKULIS, PAUL N.

REGISTRATION NUMBER: 16773

REFERENCE/DOCKET NUMBER: 91913/1107/US/ST

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 861-3000

TELEFAX: (202) 822-0944

TELEX: 6714627 CUSH

INFORMATION FOR SEQ ID NO: 59:

SEQUENCE CHARACTERISTICS:

LENGTH: 764 amino acids
 TYPE: AMINO ACID
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-741-453A-59

Query Match 12.2%; Score 403.5; DB 4; Length 764;
 Best Local Similarity 23.9%; Pred. No. 6,4e-27;
 Matches 146; Conservative 96; Mismatches 263; Indels 105; Gaps 13;

```

QY 19 PGSTHASERSOGSLIPAHPSALALASNTTASGLKLEDFQSOL-----SSIQALDL 71
DB 162 PYMTSIVNMFQGL-----CNETLTLLKLYNNMFTSVGYAFNGTKIDAVYL 207
QY 72 SWNA-IRSHPEAFSTLSLVK-LDLFDNOLFTTPLAGLGIMHLKLGKLNLAQAFSD 129
DB 208 NKNKYLIVDKDARGVYSGPSLDVSGTSTALPFSKGLHKLKLIARNTWTKRLPLSL 267
QY 130 SFPKRLLEVPYAVQCCPY-----GMCASF-----154
DB 268 SFHLTRADLSYSHCCAFKNKKRIGLIESLMCNESMOSLRKRSVNALNSPLHQEYE 327
QY 155 -----FKASGOW-----EAPDLHLDDESSKRPGLGLARQAEHNYDQ 191
DB 328 ENLGSIVGYKESKPDQTHNNAHYVFFEEQDEITIGFQGLKNPQRETLQAFDSHYDY 387
QY 192 DLDELQLEMDSRPHSVOCSPTPGPKCEYLEFESWGIRLAWAIVLVLCNGVLVLT 251
DB 388 T-----CGDSE---DMVCTPKRSDENPCEDINGYKFLRVVVFSLALLGVFVLL- 437
QY 252 VRAGGAPAPLPVYVGAIGANTLTGISGLASVDALTFGQSESGARWETGLCRAT 311
DB 438 ILTSHYKLNVPFLMCNLAFAFCMGMYLLIASVDLYTHSEYNAIDMOTGPGCNTA 497
QY 312 GFLAVIGSASVLLTLAVQCSVSVSVARAYGKSPSLGVRAGVGLCLAGLAALPL 371
DB 498 GFTTFASLSVYTLVTLERWYATTFAMRLDRKIRLRHACALIMVGWVCCFLLALLPL 557
QY 372 ASVGEYASPLCLPYAPRPGOPALGFTVALVMNSFCPLVAGATIKLYCDLPRGDFE- 430
DB 558 VGISYAVKVSICLPM--DETPLALAYIVFVLNLVAFVIVCCVYKITYITRNPOYNP 615
QY 431 AVMDCAMRHWAMLIIPADGLLYCPVALISFASMLGLEPYTPAEVKSGLVLLVPLACLNP 490
DB 616 GDKDKRIKAKRAVLIETFTICMAPISFYALSAIUNKPLITVNSKILLVLFYTLNSCANP 675
QY 491 LLYLLEFHFRRDLRLRPRAGDSGLPAAAG-----ELEKSSCDSTQALVAF 539
DB 676 FLVAIFTRAFORDVFILLSKFETCKRQAOAYRGQVRPKNSTDIQVOKVTYHDMRGGLHNK 735
QY 540 SDVDLLEAS 549
DB 736 EDVYELIENS 745

```

RESULT 14

US-07-757-342D-10
 Sequence 10, Application US/0757342D
 Patent No. 6218509

GENERAL INFORMATION:

APPLICANT: IGARASHI, Masao
 MINEGISHI, Takashi
 NAKAMURA, Kazuto
 TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS &
 CUSHMAN
 STREET: 130 Water Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: US

ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/757,342D
 FILING DATE: 10-Sep-1991
 CLASSIFICATION: <unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: BUCKLEY, Linda M.
 REGISTRATION NUMBER: 31003
 REFERENCE/DOCKET NUMBER: 41226
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617)523-3400
 TELEFAX: (617)523-6440
 TELETYPE: 200291 STRE UR
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 674 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 10:
 US-07-757-342D-10

Query Match 11.5%; Score 381; DB 4; Length 674;
 Best Local Similarity 25.3%; Pred. No. 5.1e-25;
 Matches 122; Conservative 91; Mismatches 241; Indels 28; Gaps 10;

```

QY 47 SNTTASGLKLEDFQSOL-----LQALDSWNA-IRSHPEAFSTLSLVKLDLTDN 98
DB 148 NNESTYTLKLYGFEFEVOSHANGTTLTSLKENVHLEKMHGARGATGPKTLDISS 207
QY 99 QLTPLPLAGGLMLHKLKGNLALSOAFKDSFPKRLILEVPYAVQCCPYGMCASFEXAS 158
DB 208 KLOALPSYGLSEIQRLIATSSYSCLKLPSERFVNLLENTLTPYSNCCA-----FRNL 260
QY 159 GOWEAEHLHDESSKRPGLGLARQAEH--YDQDLDELQLEMEDSK-----PHPSVOC 211
DB 261 PTKRONSHSISENFSKO-CGSTVKKVSNKTLYSMLSESLSGMDYEGFCLP-KTPRC 318
QY 212 SPTGPKPKCEYLEFESWGIRLAWAIVLVLCNGVLVLTVPAGGAPAPLPKVFYVGAIA 271
DB 319 APEPDAENPCEDINGYDPLRVLMINILALMGN-MYLFVLLTSRYKLTVPFLMCNLS 377
QY 272 GANTLTGISGLASVDALTFGQSESGARWETGLCRATGFLAVLGSSEASVLLTLAAV 331
DB 378 FADFCMLYLLILLASVDSQTKGQYNNHAIQMGTSGCCSTAGFTTFVASELSYTLTVTL 437
QY 332 QCSVSVSVARAYGKSPSLGVRAGVGLCLAGLAALPLASVGEYASPLCLPYAPREG 391
DB 438 ERWHTITVYAIHDOKLRLRAHAILMLGWLFSLLIAMLPLVGSNMYKVSICFPW--DVE 495
QY 392 QPAAIGFTVALVMNSFCPLVAGATIKLYCDLPRGDFEAV-WDCAMRHWAMLIIPADGL 450
DB 496 TTLISQYIITLILNLVAFIICACYIKIYFAVRNPMLATNKTQTKAKKAAIILFDFT 555
QY 451 LYCPVAFLSFASMLGLEPYTPAEVKSGLVLLVPLPACLNLPLLYLLEFHFRRDLRLRPR 510
DB 556 CMAPISFFAFSAFKVPLITVNSKVLVLFYTLNSCANPFLVAIFTRAFORDVFILLSK 615
QY 511 AG 512
DB 616 FG 617

```

RESULT 15

US-07-757-342D-2
 Sequence 2, Application US/0757342D
 Patent No. 6218509

GENERAL INFORMATION:

APPLICANT: IGARASHI, Masao

MINEGISHI, Takashi
NAKAMURA, Kazuo

TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
CUSHMAN

STREET: 130 Water Street

CITY: Boston

STATE: Massachusetts

COUNTRY: US

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA: US/07/757,342D

APPLICATION NUMBER: US/07/757,342D

FILING DATE: 10-Sep-1991

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: BUCKLEY, Linda M.

REGISTRATION NUMBER: 31003

REFERENCE/DOCKET NUMBER: 41226

TELEPHONE: (617)523-3400

TELEFAX: (617)523-6440

TELEX: 200291 STRE UR

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 699 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-07-757-342D-2

Query Match 11.5%; Score 381; DB 4; Length 699;
Best Local Similarity 25.3%; Pred. No. 5.4e-25;

Matches 122; Conservative 91; Mismatches 241; Indels 28; Gaps 10;

Db	581	CMAPIFFAISAFAKVPILTVNSKVLVLYFPINSCANPFLIATFTKTFORDPFLLSK	640
QY	511	AG 512	
Db	641	FG 642	

Search completed: November 8, 2002, 19:37:23
Job time: 16.9467 secs

47 SNTTASGKLEXTFSQSS-----LQALDLSMNA--IRSIHPEAFSTLHSLVKLIDLTDN 98
173 NNEGVTLKLVNGFEVQSHAFNGTTLTSLKENVHLEKMHNGAFRGATGPKTLDISST 232
99 QLTTLPLAGLGLMHLKLGKGLALSOAFSKDPSFKRLIEVPYAOCCPYGCASFRRAS 158
233 KLOALPSYGLSIORLITATSSYSLKLPSTRETFVNLLEATILTPSHCCA-----FNNL 285
159 GQWEAEDLHLDDESSKRPGLILARQAEH--YQDDLELOLEMEDSK-----PHPSVQC 211
286 PTKQGNFHSISENSKQ--CESTVRKVSNTKLYSSMLAESELGWDYEGCLP--KTPRC 343
212 SPTGPFPKCEYLPESMGIRLAVNAIVLVCNGLVLLTVFAGGAPPLPYKFFVGAIA 271
344 APEPDAPNCEIDIMGYDPLRVILMLITLALGN--WTVLFVLLTSRKILTVPRFLMCNLS 402
272 GANTLTGISCGLASVDALTFQFSEYGARWETGLGCRATGFLAVLGSFASVLLTLAAV 331
403 FADFCMGILYLLIAVDSQTKGYNNHAIIDMOTGSGCSTAGFFTFASLSYTLTVITL 462
332 QCSVSVSCYRAVGKSPSGSVAGVLCGLALAGLAAALPLASVGEYGASPLCLPYAPREG 391
463 ERMWHTTYAIHLDOKLRHAILIMLGMFLSSLIAMDPLVGVSNMYKYSICPPM--DYE 520
392 QPAALGFVALVMMNSFCFLVAVGAYIKLYCDLPRGDFEAV--WDCAMVRHVAWLIFADGL 450
521 TILSQVYIITLILINVAFFIICACYIKIYPAVRNPELMATNMDTKIAKMAILLIFDFT 580
451 LVCPAVAFISFASMLGLFPYTPPEAKSVLLVLPACLNPLLYLLFNPFRDRLRLRPR 510

November 8, 2002, 19:31:14 ; Search time 32.8291 Seconds
(without alignments)
2569.297 Million cell updates/sec

Title:	US-09-851-595-5
Perfect score:	3307
Sequence:	1 NTHYRESWYACRYRSGIPG.....GGLSGGGCFQPSGLAFASHV 633

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

```

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

```

Post-processing:	Minimum Match	0%
	Maximum Match	100%

Listing first 45 summaries

Database

```

1: /SIDS2/gcgdata/genseq/genseqp -emb1 /AA1980.DAT.*
2: /SIDS2/gcgdata/genseq/genseqp -emb1 /AA1981.DAT.*
3: /SIDS2/gcgdata/genseq/genseqp -emb1 /AA1982.DAT.*
4: /SIDS2/gcgdata/genseq/genseqp -emb1 /AA1983.DAT.*
5: /SIDS2/gcgdata/genseq/genseqp -emb1 /AA1984.DAT.*
6: /SIDS2/gcgdata/genseq/genseqp -emb1 /AA1985.DAT.*
7: /SIDS2/gcgdata/genseq/genseqp -emb1 /AA1986.DAT.*
8: /SIDS2/gcgdata/genseq/genseqp -emb1 /AA1987.DAT.*
9: /SIDS2/gcgdata/genseq/genseqp -emb1 /AA1988.DAT.*
10: /SIDS2/gcgdata/genseq/genseqp -emb1 /AA1989.DAT.*
11: /SIDS2/gcgdata/genseq/genseqp -emb1 /AA1990.DAT.*
12: /SIDS2/gcgdata/genseq/genseqp -emb1 /AA1991.DAT.*
13: /SIDS2/gcgdata/genseq/genseqp -emb1 /AA1992.DAT.*
14: /SIDS2/gcgdata/genseq/genseqp -emb1 /AA1993.DAT.*
15: /SIDS2/gcgdata/genseq/genseqp -emb1 /AA1994.DAT.*
16: /SIDS2/gcgdata/genseq/genseqp -emb1 /AA1995.DAT.*
17: /SIDS2/gcgdata/genseq/genseqp -emb1 /AA1996.DAT.*
18: /SIDS2/gcgdata/genseq/genseqp -emb1 /AA1997.DAT.*
19: /SIDS2/gcgdata/genseq/genseqp -emb1 /AA1998.DAT.*
20: /SIDS2/gcgdata/genseq/genseqp -emb1 /AA1999.DAT.*
21: /SIDS2/gcgdata/genseq/genseqp -emb1 /AA2000.DAT.*
22: /SIDS2/gcgdata/genseq/genseqp -emb1 /AA2001.DAT.*
23: /SIDS2/gcgdata/genseq/genseqp -emb1 /AA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3305	99.9	633	23	AA666139	Human LGR6 polypep
2	3015	91.2	736	23	AA666140	Human LGR6 polypep
3	3015	91.2	948	23	AAE23418	Human G-protein co
4	3015	91.2	967	23	AA666141	Human LGR6 polypep
5	2860	89.5	847	20	AAV53575	Human gonadotropin
6	2960	89.5	940	23	AAU79941	Human orphan GPCR
7	2637	79.7	967	23	AA666138	Mouse LGR6 polypep
8	2494.5	75.4	691	22	ABG0314	Novel human diagno
9	2450	74.1	497	22	AAV53574	Human gonadotropin
10	2033	61.5	395	22	AA667556	A human guanosine

11	1275	38.6	907	20	AA938689	Human Hc38 protein
12	1275	38.6	907	21	AA930682	Human G protein-co
13	1273	38.5	693	20	AA142169	Human LGF5 protein
14	1269	38.4	907	21	AA190867	Human mutant G pro
15	1206	36.5	230	22	AA950066	Human protein sequ
16	969	29.3	951	20	AA1442168	Human LGF4 protein
17	967	29.2	951	23	AA805224	Rat LGR4-like GPCR
18	958	29.0	634	20	AA862318	Human 7-transmembr
19	957	28.9	631	20	AA133767	Human AXOR33 prote
20	957	28.9	951	22	AA828835	Human REC4P polype
21	957	28.9	951	22	AA868875	Human AOMF05 prote
22	952	28.8	949	20	AA933904	Human AOMF05 prote
23	952	28.8	951	20	AA933905	Human AOMF05 prote
24	952	28.8	951	20	AA933906	Human AOMF05 prote
25	952	28.8	951	22	AAU32972	Novel human secret
26	952	28.8	951	22	AAU32972	Human gonadotropin
27	901	27.2	171	20	AA153571	Human gonadotropin
28	858	25.9	165	23	ABP41438	Human ovularin anti
29	745.5	22.5	692	23	AAW93890	Human HG38 protein
30	639.5	19.3	282	23	ABP40000	Human ovularin anti
31	608.5	18.4	1300	22	AB664083	D. melanogaster pe
32	539	16.3	861	22	AA886961	Human gonadotropin
33	502.5	15.2	202	20	AA153572	Human gonadotropin
34	474	14.3	220	20	AA153578	Human gonadotropin
35	474	14.3	254	20	AA153578	Human gonadotropin
36	424	12.8	696	14	AA830513	N-terminal of LH r
37	424	12.8	696	14	AA830521	N-terminal of LH r
38	424	12.8	703	14	AA830507	N-terminal of LH r
39	418	12.6	636	14	AA830519	N-terminal of LH r
40	417	12.6	650	14	AA830514	N-terminal of LH r
41	417	12.6	656	14	AA830526	N-terminal of LH r
42	417	12.6	697	14	AA830508	N-terminal of LH r
43	416.5	12.6	699	14	AA830515	N-terminal of LH r
44	416.5	12.6	764	12	AA112504	Canine thyroid st
45	415	12.5	620	14	AA830522	N-terminal of LH r

ALIGNMENTS

RESULT 1	.
AAG66139	
ID	AAG66139 standard; Protein; 633 AA

AC AAG66139;

DT 13-MAR-2002 (first entry)

aa	Human LGR6 polypeptide (clone fahr)
DE	

XX Large G-protein coupled receptor 6; LGR6; GPCR; immunomodulator;
KW neuroactive ligand-receptor interaction; antiParkinsonian;

KW anti-HIV; cytostatic; nootropic; neuroprotective; antidepressant; tranquilizer; antimigraine
KW anticonvulsant; neuroleptic; antidepressant; vasodilator; hypotensive; cardiant

osteopathic; anorectic; gene therapy; human antiarrhythmic; anorectic; gene therapy; human

05 Homo sapiens

AA FH	Key	Location/Qualifiers

FT	Misc-difference	57
	[label] = unknown	

```

11      /note= "encoded by CTN
12      FT

```

PN W0200185768-A2

PD 15-NOV-2001

08-MAY-2001; 2001WO-US15002-
PF

08-MAY-2000; 2000US-0566588.

PA (MILL-) MILLENNIUM PHARM INC

PI Gu W;
 XX WPI: 2002-055584/07.
 DR N-PSDB; AA167922, AA167923.

PT Novel isolated large G-protein coupled receptor 6 polypeptide, useful
 PT for diagnosing and treating weight disorder, metabolic disorder,
 PT central nervous system disorder, endocrine disorder and cardiovascular
 PT disorder

XX Example 1; Fig 5; 198pp; English.

CC The invention provides isolated large G-protein coupled receptor 6 (LGR6)
 CC polypeptides and polynucleotides encoding them. The LGR6 polypeptides and
 CC polynucleotides are useful as targets for developing modulating agents
 CC that regulate a variety of cellular process, e.g. neural and endocrine
 CC processes, as well as thermogenesis. They are useful for developing novel
 CC diagnostic and therapeutic agents for LGR6 associated disorders such as
 CC weight disorders (anorexia, obesity), eating disorders, acquired
 CC immunodeficiency syndrome (AIDS)-related wasting, cancer-related wasting,
 CC metabolic disorders, central nervous system disorders (e.g. Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease, Gilles de la
 CC Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis,
 CC epilepsy, Jakob-Creutzfeldt disease, depression, schizophrenia,
 CC obsessive-compulsive disorder, anxiety, panic disorder, migraine),
 CC endocrine disorders (hypo and hyperthyroidism, dwarfism, gigantism),
 CC acromegaly), and cardiovascular disorders (atherosclerosis, restenosis,
 CC hypertension, myocardial infarction, arrhythmia). The LGR6 polypeptides,
 CC polynucleotides and antibodies are useful in screening assays, detection
 CC assays (e.g., chromosomal mapping, tissue typing, forensic biology),
 CC predictive medicine (e.g. diagnostic assays, prognostic assays,
 CC monitoring clinical trials and pharmacogenomics), and in methods of
 CC treatment (e.g. therapeutic and pharmacologic). The present sequence
 CC represents a human LGR6 polypeptide.

Sequence 633 AA:

Query Match 99.9%; Score 3305; DB 23; Length 633;
 Best Local Similarity 100.0%; Pred. No. 2, 66-293;
 Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTHYRESWYACRYSRGSTHSAVERSOGLSLPAHPSALAAASNTTASGKLEXTF 60
 DB 1 NTHYRESWYACRYSRGSTHSAVERSOGLSLPAHPSALAAASNTTASGKLEXTF 60
 QY 61 SOLSLQALDLSMAIRSIHPEAFSTLSLVKLDLTNQLTTLPLAGLGLMHLKGNL 120
 DB 61 SOLSLQALDLSMAIRSIHPEAFSTLSLVKLDLTNQLTTLPLAGLGLMHLKGNL 120
 QY 121 ALSQAFKSDSPKRILEVPAVYOCPCYGMCAFFKASGOWEADLHLDDESSKRPGL 180
 DB 121 ALSQAFKSDSPKRILEVPAVYOCPCYGMCAFFKASGOWEADLHLDDESSKRPGL 180
 QY 181 LARQENHYDDDELDELQLEMEDSKRHPVOCSPNPPGPKPEYTFESGIRLAWAIVLL 240
 DB 181 LARQENHYDDDELDELQLEMEDSKRHPVOCSPNPPGPKPEYTFESGIRLAWAIVLL 240
 QY 241 SYLCNGVYLTVRAGGAPLPVYKFFVGAAGANTLGLISGGLASVDALTFQGFSEGA 300
 DB 241 SYLCNGVYLTVRAGGAPLPVYKFFVGAAGANTLGLISGGLASVDALTFQGFSEGA 300
 QY 301 RWETGLGCRATGFLAVLGSEASVLLTLTAVVOCSSVSCVRANGKSPSLGVRAGVLCGL 360
 DB 301 RWETGLGCRATGFLAVLGSEASVLLTLTAVVOCSSVSCVRANGKSPSLGVRAGVLCGL 360
 QY 361 ALAGLAAALPLASVGEYGAAPLCPYAPPEGPAALGFTYALVMNMSFCLVAVAGAYIKL 420
 DB 361 ALAGLAAALPLASVGEYGAAPLCPYAPPEGPAALGFTYALVMNMSFCLVAVAGAYIKL 420
 QY 421 YCDLPGRGDEAVWDCAMVHVAMLLIFADGLTCPPVAFSLFASMLGLFPVTPEAVKSVLLV 480
 DB 421 YCDLPGRGDEAVWDCAMVHVAMLLIFADGLTCPPVAFSLFASMLGLFPVTPEAVKSVLLV 480

QY 481 VLPLPACLNPLLYLLFNPNFRDDLRRLRPAGDSGELAAVAAAGELKSSCDSTQALVARS 540
 DB 481 VLPLPACLNPLLYLLFNPNFRDDLRRLRPAGDSGELAAVAAAGELKSSCDSTQALVARS 540
 QY 541 DVDLLILEASEAGRPGLERYGFPSTVLLISCOOPGAPRLGSGHCVEEGHFGNPOPSMDG 600
 DB 541 DVDLLILEASEAGRPGLERYGFPSTVLLISCOOPGAPRLGSGHCVEEGHFGNPOPSMDG 600
 QY 601 ELLLRAGSTPAGGGLSGGGGFPSPGLAFASHV 633
 DB 601 ELLLRAGSTPAGGGLSGGGGFPSPGLAFASHV 633

RESULT 2
 ID AAG66140 standard; Protein: 736 AA.
 AC AAG66140;

DT 13-MAR-2002 (first entry)
 DE Human LGR6 polypeptide (clone fahr).

KW Large G-protein coupled receptor 6; LGR6; GPCR; immunomodulator;
 KW anti-HIV; cyostatic; nootropic; neuroprotective; antiparkinsonian;
 KW anticonvulsant; neuroleptic; antidepressant; tranquilizer; antimigraine;
 KW osteopathic; antiarteriosclerotic; vasotropic; hypotensive; cardiant;
 KW antiarrhythmic; anorectic; gene therapy; human.

OS Homo sapiens.

PN WO200185766-A2.

PD 15-NOV-2001.

PF 08-MAY-2001; 2001WO-US15002.

PR 08-MAY-2000; 2000US-0566588.

PA (MILL-) MILLENNIUM PHARM INC.

PI Gu W;

DR WPI: 2002-055584/07.

DR N-PSDB; AA167924, AA167925.

PT Novel isolated large G-protein coupled receptor 6 polypeptide, useful
 PT for diagnosing and treating weight disorder, metabolic disorder,
 PT central nervous system disorder, endocrine disorder and cardiovascular
 PT disorder

Claim 9; Fig 8; 198pp; English.

CC The invention provides isolated large G-protein coupled receptor 6 (LGR6)
 CC polypeptides and polynucleotides encoding them. The LGR6 polypeptides and
 CC polynucleotides are useful as targets for developing modulating agents
 CC that regulate a variety of cellular process, e.g. neural and endocrine
 CC processes, as well as thermogenesis. They are useful for developing novel
 CC diagnostic and therapeutic agents for LGR6 associated disorders such as
 CC weight disorders (anorexia, obesity), eating disorders, acquired
 CC immunodeficiency syndrome (AIDS)-related wasting, cancer-related wasting,
 CC metabolic disorders, central nervous system disorders (e.g. Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease, Gilles de la
 CC Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis,
 CC epilepsy, Jakob-Creutzfeldt disease, depression, schizophrenia,
 CC obsessive-compulsive disorder, anxiety, panic disorder, migraine),
 CC endocrine disorders (hypo and hyperthyroidism, dwarfism, gigantism),
 CC acromegaly), and cardiovascular disorders (atherosclerosis, restenosis,
 CC hypertension, myocardial infarction, arrhythmia). The LGR6 polypeptides,
 CC polynucleotides and antibodies are useful in screening assays, detection
 CC assays (e.g., chromosomal mapping, tissue typing, forensic biology),
 CC predictive medicine (e.g. diagnostic assays, prognostic assays,
 CC monitoring clinical trials and pharmacogenomics), and in methods of

293 294 TTTT TTTT SVT CNGT YI TVEAGGPAPLPVKFVVGAIAGANTLTGISCGLLASVDALTFG 293

293 294 TTTT TTTT SVT CNGT YI TVEAGGPAPLPVKFVVGAIAGANTLTGISCGLLASVDALTFG 293

```

Db 549 VVAIVLTVLNCNGVLLTVFAGGPAFLPVKTVGAIAGANTLTGISCGLASVDALITFG 608
QY 294 OFSEYGARWETLGRATGFLAVLGSSEASVLLTLTAAVOCSSVSCVRAVYKSPSLGSVR 353
Db 609 OFSEYGARWETLGRATGFLAVLGSSEASVLLTLTAAVOCSSVSCVRAVYKSPSLGSVR 668
QY 354 AGVLCIALAGLAALPLASVGEYASPLCLPYAPREGOPALGFTVALVMNSFCFLVY 413
Db 669 AGVLCIALAGLAALPLASVGEYASPLCLPYAPREGOPALGFTVALVMNSFCFLVY 728
QY 414 AGATIKLYCDLPGRDFEAVWDCAMVRHVAWLIFADGLLYCPVAFLSFASMLGLFPVTEPA 473
Db 729 AGATIKLYCDLPGRDFEAVWDCAMVRHVAWLIFADGLLYCPVAFLSFASMLGLFPVTEPA 788
QY 474 VKSVLLVVLPLPACLNPLLYLTFNPHFRDDLRRLRPRAAGSGPLAAYAAAGLEKSSCDST 533
Db 789 VKSVLLVVLPLPACLNPLLYLTFNPHFRDDLRRLRPRAAGSGPLAAYAAAGLEKSSCDST 848
QY 534 QALVAFSDVDLILEASERGPGLGTYGFPSTVLISCOQPGAPRLGSHCVPEPGNHNFGN 593
Db 849 QALVAFSDVDLILEASERGPGLGTYGFPSTVLISCOQPGAPRLGSHCVPEPGNHNFGN 908
QY 594 POPSMDEGLLIRAGSTPAGGSLGGGFGPSSGLAFASHV 633
Db 909 POPSMDEGLLIRAGSTPAGGSLGGGFGPSSGLAFASHV 948

RESULT 4
AAG66141
ID AAG66141 standard; Protein; 967 AA.
AC AAG66141;
AS 13-MAR-2002 (first entry)
DE Human LGR6 polypeptide (clone Fb150881).
KM Large G-protein coupled receptor 6; LGR6; GPCR; immunomodulator;
KW anti-HIV; cytostatic; nootropic; neuroprotective; antiparkinsonian;
KW anticonvulsant; neuroleptic; antidepressant; tranquilizer; antidiarrheic;
KW osteopathic; antiatherosclerotic; vasodilator; hypotensive; cardiac;
KW antirhythmic; anorectic; gene therapy; human.
OS Homo sapiens.
PN WO200185768-A2.
PD 15-NOV-2001.
PE 08-MAY-2001; 2001WO-US15002.
PR 08-MAY-2000; 2000US-0566588.
PA (MILL-) MILLENNIUM PHARM INC.
PI Gu W;
PS WPI: 2002-055584/07.
DR N-PSDB; AAI67926, AAI67927.
PT Novel isolated large G-protein coupled receptor 6 polypeptide, useful
PT for diagnosing and treating weight disorder, metabolic disorder,
PT central nervous system disorder, endocrine disorder and cardiovascular
PT disorder.
PS Claim 9; Fig 15; 198pp; English.
CC The invention provides isolated large G-protein coupled receptor 6 (LGR6)
CC polypeptides and polynucleotides encoding them. The LGR6 polypeptides and
CC polynucleotides are useful as targets for developing modulating agents and
CC processes, as well as thermogenesis. They are useful for developing novel

```

```

CC diagnostic and therapeutic agents for LGR6 associated disorders such as
CC weight disorders (anorexia, obesity), eating disorders, acquired
CC immunodeficiency syndrome (AIDS)-related wasting, cancer-related wasting,
CC metabolic disorders, central nervous system disorders (e.g. Alzheimer's
CC disease, Parkinson's disease, Huntington's disease, Gilles de la
CC Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis,
CC obsessive-compulsive disorder, anxiety, panic disorder, migraine,
CC endocrine disorders (hypo and hyperthyroidism, dwarfism, gigantism,
CC acromegaly), and cardiovascular disorders (atherosclerosis, restenosis,
CC polynucleotides and antibodies are useful in screening assays, detection
CC assays (e.g., chromosomal mapping, tissue typing, forensic biology),
CC predictive medicine (e.g. diagnostic assays, prognostic assays,
CC monitoring clinical trials and pharmacogenomics), and in methods of
CC treatment (e.g. therapeutic and prophylactic). The present sequence
CC represents a human LGR6 polypeptide.
SQ Sequence 967 AA;
Query Match 91.2%; Score 3015; DB 23; Length 967;
Best Local Similarity 99.3%; Pred. No. 1.6e-266;
Matches 576; Conservative 2; Mismatches 2; Indels 0; Caps 0;
QY 54 KLENDTFSQSSLDALDSWNAIRSIHPEAFSTLSLVKLDITDQNTTPIAGLGIMH 113
Db 388 ELGADTFSSQLSLQALDLSWNAIRSIHPEAFSTLSLVKLDITDQNTTPIAGLGIMH 447
QY 114 LKLGNLALSQAFSKDSEFPKRLILEVYAYOCGCGKCAFFRASGWEADLHDEES 173
Db 448 LKLGNLALSQAFSKDSEFPKRLILEVYAYOCGCGKCAFFRASGWEADLHDEES 507
QY 174 SKRPLGILAAQANHHYDQIDDELQLEMEDSKPHPSVQCSPTPEPFKCELYFESWGRILA 233
Db 508 SKRPLGILAAQANHHYDQIDDELQLEMEDSKPHPSVQCSPTPEPFKCELYFESWGRILA 567
QY 234 VVAIVLTVLNCNGVLLTVFAGGPAFLPVKTVGAIAGANTLTGISCGLASVDALITFG 293
Db 568 VVAIVLTVLNCNGVLLTVFAGGPAFLPVKTVGAIAGANTLTGISCGLASVDALITFG 627
QY 294 OFSEYGARWETLGRATGFLAVLGSSEASVLLTLTAAVOCSSVSCVRAVYKSPSLGSVR 353
Db 628 OFSEYGARWETLGRATGFLAVLGSSEASVLLTLTAAVOCSSVSCVRAVYKSPSLGSVR 687
QY 354 AGVLCIALAGLAALPLASVGEYASPLCLPYAPREGOPALGFTVALVMNSFCFLVY 413
Db 688 AGVLCIALAGLAALPLASVGEYASPLCLPYAPREGOPALGFTVALVMNSFCFLVY 747
QY 414 AGATIKLYCDLPGRDFEAVWDCAMVRHVAWLIFADGLLYCPVAFLSFASMLGLFPVTEPA 473
Db 748 AGATIKLYCDLPGRDFEAVWDCAMVRHVAWLIFADGLLYCPVAFLSFASMLGLFPVTEPA 807
QY 474 VKSVLLVVLPLPACLNPLLYLTFNPHFRDDLRRLRPRAAGSGPLAAYAAAGLEKSSCDST 533
Db 808 VKSVLLVVLPLPACLNPLLYLTFNPHFRDDLRRLRPRAAGSGPLAAYAAAGLEKSSCDST 867
QY 534 QALVAFSDVDLILEASERGPGLGTYGFPSTVLISCOQPGAPRLGSHCVPEPGNHNFGN 593
Db 868 QALVAFSDVDLILEASERGPGLGTYGFPSTVLISCOQPGAPRLGSHCVPEPGNHNFGN 927
QY 594 POPSMDEGLLIRAGSTPAGGSLGGGFGPSSGLAFASHV 633
Db 928 POPSMDEGLLIRAGSTPAGGSLGGGFGPSSGLAFASHV 967

RESULT 5
AAV53575
ID AAV53575 standard; Protein; 847 AA.
AC AAV53575;
AS 15-FEB-2000 (first entry)
DT
XX

```

Db	609	AGATIKLYCDLPKGDGEAAYWCAMRYHAWMLIFADGLLYCPVAFLSFSMGLFPVTP	668
QY	474	VKSYLVLPVLPACINPELLYLLENHFRNDLRRLPRAGDSGPLAYAAAGLEKSCDST	533
Db	669	VKSYLVLPVLPACINPELLYLLENHFRNDLRRLPRAGDSGPLAYAAAGLEKSCDST	728
QY	534	QALVAFSVDVLLILEASAGRPGLTEYGFPSYTLISCOQPGAPRLGSHCYEPPENHFN	593
Db	729	QALVAFSVDVLLILEASAGRPGLTEYGFPSYTLISCOQPGAPRLGSHCYEPPENHFN	788
QY	594	POPMDGELLIRAGSFPAGGSLSGGGGFP	624
Db	789	POPMDGELLIRAGSFPAGGSLSGGGGFP	819

QY 114 LKLGNTALSOAFSKDSKDFPKRLILEVPAAYOCCPYGMKASFFKASQGWAEADLHLDDEES 173
 DB 402 LKLGNTALSOAFSKDSKDFPKRLILEVPAAYOCCPYGMKASFFKASQGWAEADLHLDDEES 461
 QY 174 SKRPLGLLARAQENHYDODLDELQLEMEDESKRPSVOCSPPGPKFCETLFEWGIKRLA 233
 DB 462 SKRPLGLLARAQENHYDODLDELQLEMEDESKRPSVOCSPPGPKFCETLFEWGIKRLA 521
 QY 234 VMAIVLLSVLCNGILVLTVEFAGGPAPLPYKFWVGAAGANTLTGISGGLASVDALTFEG 293
 DB 522 VMAIVLLSVLCNGILVLTVEFAGGPAPLPYKFWVGAAGANTLTGISGGLASVDALTFEG 581
 QY 294 QFSEYGARWETGIGCRATGGLAYLVSSEASVLLTLTAAVOCSSVSCVRAVYKSPSLGSVR 353
 DB 582 QFSEYGARWETGIGCRATGGLAYLVSSEASVLLTLTAAVOCSSVSCVRAVYKSPSLGSVR 641
 QY 354 AGVGLGCLALAGLAAALPLASVGEYASPLCLPYAPPEGQPALGFTVALVMNNSFCFLV 413
 DB 642 AGVGLGCLALAGLAAALPLASVGEYASPLCLPYAPPEGQPALGFTVALVMNNSFCFLV 701
 QY 414 AGAYIKLYCDLPRGDFEAVWDCAMVRHVAMLIADGLLYCPVAFISFASMLGLFVTPPEA 473
 DB 702 AGAYIKLYCDLPRGDFEAVWDCAMVRHVAMLIADGLLYCPVAFISFASMLGLFVTPPEA 761
 QY 474 VKSVLVLPPLPACLNPLLYLLEFNPFRDRLRLRPAGDSGLATYAAAGLEKSSCDST 533
 DB 762 VKSVLVLPPLPACLNPLLYLLEFNPFRDRLRLRPAGDSGLATYAAAGLEKSSCDST 821
 QY 534 QALVAFSDVDLILEASEAGRPGLTETGFPSTVLTISCCQPGAPRLKESHCVEPBGNNHFGN 593
 DB 822 QALVAFSDVDLILEASEAGRPGLTETGFPSTVLTISCCQPGAPRLKESHCVEPBGNNHFGN 881
 QY 594 POPMDELLLRAGSTIPAGGGLSGGGFOP 624
 DB 882 POPMDELLLRAGSTIPAGGGLSGGGFOP 912

RESULT 7
 AAG66138
 ID AAG66138 standard; Protein: 967 AA.
 AC AAG66138;
 DT 13-MAR-2002 (first entry)
 DE Mouse LGR6 polypeptide (clone flmzb048h10).
 KW Large G-protein coupled receptor 6; LGR6; GPCR; Immunomodulator;
 KW anti-HIV; cytoskeletal; nootropic; neuroprotective; antiparkinsonian;
 KW anticonvulsant; neuroleptic; antidepressant; tranquilizer; antimigraine;
 KW osteopathic; antiarteriosclerotic; vasodilator; hypotensive; cardiant;
 KW antiarhythmic; anorectic; gene therapy; mouse.
 OS Mus musculus.
 XX
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..23
 FT Protein /note= "signal peptide"
 FT /note= "24..967
 FT /note= "mature protein"
 PN W0200185768-A2.
 PD 15-NOV-2001.
 XX 08-MAY-2001; 2001WO-US15002.
 PF 08-MAY-2000; 2000US-0566588.
 PR (MILL-) MILLENNIUM PHARM INC.
 PA
 XX Gu W;
 XX

DR WPI: 2002-055584/07.
 DR N-PSDB: AAI67920, AAI67921.
 XX
 PT Novel isolated large G-protein coupled receptor 6 polypeptide, useful
 PT for diagnosing and treating weight disorder, metabolic disorder,
 PT central nervous system disorder, endocrine disorder and cardiovascular
 PT disorder -
 XX
 PS Example 1; Fig 1; 198pp; English.
 XX
 CC The invention provides isolated large G-protein coupled receptor 6 (LGR6)
 CC polypeptides and polynucleotides encoding them. The LGR6 polypeptides and
 CC polynucleotides are useful as targets for developing modulating agents
 CC that regulate a variety of cellular processes, e.g. neural and endocrine
 CC processes, as well as thermogenesis. They are useful for developing novel
 CC diagnostic and therapeutic agents for LGR6 associated disorders such as
 CC weight disorders (anorexia, obesity), eating disorders, acquired
 CC immunodeficiency syndrome (AIDS)-related wasting, cancer-related
 CC metabolic disorders, central nervous system disorders (e.g. Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease, Gilles de la
 CC Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis,
 CC obsessive-compulsive disorder, anxiety, panic disorder, schizophrenia,
 CC endocrine disorders (hypo and hyperthyroidism, dwarfism, gigantism,
 CC acromegaly), and cardiovascular disorders (atherosclerosis, restenosis,
 CC polynucleotides and antibodies are useful in screening assays, detection
 CC assays (e.g., chromosomal mapping, tissue typing, forensic biology),
 CC predictive medicine (e.g. diagnostic assays, prognostic assays,
 CC monitoring clinical trials and pharmacogenomics), and in methods of
 CC treatment (e.g. therapeutic and prophylactic). The present sequence
 CC represents a mouse LGR6 polypeptide.
 CC
 SQ Sequence 967 AA;
 Query Match 79.7%; Score 2637; DB 23; Length 967;
 Best Local Similarity 87.2%; Pred. No. 6.3e-232;
 Matches 506; Conservative 28; Mismatches 46; Indels 0; Gaps 0;
 QY 54 LKXDTFSQSSLOADLISWNAIRSIHPEAFSTLSLVKLDLTDNOLTTPLAGLGLMH 113
 DB 388 EIGADYFSQSSLOADLISWNAIRSIHPEAFSTLSLVKLDLTDNOLTTPLAGLGLMH 113
 QY 114 LKLGNTALSOAFSKDSKDFPKRLILEVPAAYOCCPYGMKASFFKASQGWAEADLHLDDEES 173
 DB 448 LKLGNTALSOAFSKDSKDFPKRLILEVPAAYOCCPYGMKASFFKASQGWAEADLHLDDEES 507
 QY 174 SKRPLGLLARAQENHYDODLDELQLEMEDESKRPSVOCSPPGPKFCETLFEWGIKRLA 233
 DB 508 KRPPLGLLARAQENHYDODLDELQLEMEDESKRPSVOCSPPGPKFCETLFEWGIKRLA 567
 QY 234 VMAIVLLSVLCNGILVLTVEFAGGPAPLPYKFWVGAAGANTLTGISGGLASVDALTFEG 293
 DB 568 VMAIVLLSVLCNGILVLTVEFAGGPAPLPYKFWVGAAGANTLTGISGGLASVDALTFEG 627
 QY 294 QFSEYGARWETGIGCRATGGLAYLVSSEASVLLTLTAAVOCSSVSCVRAVYKSPSLGSVR 353
 DB 628 QFSEYGARWETGIGCRATGGLAYLVSSEASVLLTLTAAVOCSSVSCVRAVYKSPSLGSVR 687
 QY 354 AGVGLGCLALAGLAAALPLASVGEYASPLCLPYAPPEGQPALGFTVALVMNNSFCFLV 413
 DB 688 AGVGLGCLALAGLAAALPLASVGEYASPLCLPYAPPEGQPALGFTVALVMNNSFCFLV 747
 QY 414 AGAYIKLYCDLPRGDFEAVWDCAMVRHVAMLIADGLLYCPVAFISFASMLGLFVTPPEA 473
 DB 748 AGAYIKLYCDLPRGDFEAVWDCAMVRHVAMLIADGLLYCPVAFISFASMLGLFVTPPEA 807
 QY 474 VKSVLVLPPLPACLNPLLYLLEFNPFRDRLRLRPAGDSGLATYAAAGLEKSSCDST 533
 DB 808 VKSVLVLPPLPACLNPLLYLLEFNPFRDRLRLRPAGDSGLATYAAAGLEKSSCDST 867
 QY 534 QALVAFSDVDLILEASEAGRPGLTETGFPSTVLTISCCQPGAPRLKESHCVEPBGNNHFGN 593
 DB 867 QALVAFSDVDLILEASEAGRPGLTETGFPSTVLTISCCQPGAPRLKESHCVEPBGNNHFGN 912

DR N-PSDB; AAZ40460.

... Van Der Spek PJ, Heikoop JC;
PI
XX

PI van der Spek PJ, Heinooy CC,
XX
DR WPI; 1999-563673/48.
DR N-PSDB; AA240460.

XX New 7 transmembrane gonadotropin receptors, useful for screening for
PT hormone analogs and drugs -
PS Claim 1; Page 18-21; 38pp; English.

CC Sequences AAY5371-Y5378 represent fragments of 3 novel human
CC gonadotropin receptors. The novel gonadotropin receptors can be used to
CC screen for ligands of the receptors. This screen may be used in
CC biomedical and biochemical research to develop new drugs targeted to the
CC gonadotropin receptors. For example, hormone analogs which activate or
CC inhibit the function of the gonadotropin receptors or classical
CC gonadotropin receptors may be detected. The altered expression or
CC dysfunction of the gonadotropin receptors causes conditions in
CC reproductive tissues. The ligands can be used for the treatment of
CC infertility or for contraception.

Sequence 497 AA;

Query Match Best Local Similarity 74.1%; Score 2450; DB 20; Length 497;
Matches 467; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 156 KASGQWEAEDLHDDDESSKRLPLGLARQAEHNDLDELQLEMENSKPHSVQCSPTP 215
DB - 1 KASGQWEAEDLHDDDESSKRLPLGLARQAEHNDLDELQLEMENSKPHSVQCSPTP 60
QY 216 GPPKPEEYLFESMGIRLAWAIYLLSVLNGVLLTVFAGGPAFLPPVKFVGAAGANT 275
DB 61 GPPKPEEYLFESMGIRLAWAIYLLSVLNGVLLTVFAGGPAFLPPVKFVGAAGANT 120
QY 276 LNGISGLASVDALTFGGFSEYGARWETGLGCRATGFLAVLGEASVLLTLAAVQCSV 335
DB 121 LNGISGLASVDALTFGGFSEYGARWETGLGCRATGFLAVLGEASVLLTLAAVQCSV 180
QY 336 SVSCVRANGKSPSLGSRVRAVGLCLALAGLAALPLASVEYASPLCLPYAPPEGPA 395
DB 181 SVSCVRANGKSPSLGSRVRAVGLCLALAGLAALPLASVEYASPLCLPYAPPEGPA 240
QY 396 LGFTVALVWMSFCLVVAAGATIKLYCDLPRGDFEAVWDCAMVRHVAWLTFADGLLYCPV 455
DB 241 LGFTVALVWMSFCLVVAAGATIKLYCDLPRGDFEAVWDCAMVRHVAWLTFADGLLYCPV 300
QY 456 AFLFSASMLGLFPVTPPEAVKSVLLVPLPACINPLLYLLENFHRDDLRLRPRAAGDSG 515
DB 301 AFLFSASMLGLFPVTPPEAVKSVLLVPLPACINPLLYLLENFHRDDLRLRPRAAGDSG 360
QY 516 PLAYAAAGELKSSCDSTQALVAFSDVDLILEASEGRPPGLETYGFPSTVLLISCOOPGA 575
DB 361 PLAYAAAGELKSSCDSTQALVAFSDVDLILEASEGRPPGLETYGFPSTVLLISCOOPGA 420
QY 576 PRLGSHCVPEEGNHFGNPPSMOGEILLRAGSTPAGGGLSGGGGFP 624
DB 421 PRLGSHCVPEEGNHFGNPPSMOGEILLRAGSTPAGGGLSGGGGFP 469

RESULT 10

ID AAG67556 standard; Protein; 395 AA.

XX AAG67556;

XX 26-NOV-2001 (first entry)

XX A human guanosine triphosphate binding protein coupled receptor.

XX Human; guanosine triphosphate binding protein coupled receptor;

XX G protein coupled receptor; brain disease; cancer.

OS Homo sapiens.

XX WO200109323-A1.

PD 08-FEB-2001.

XX 28-JUL-2000; 2000WO-JP05070.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 18-OCT-1999; 99JP-0159590.

XX 11-JAN-2000; 2000JP-0118776.

XX 17-FEB-2000; 2000US-0183322.

XX 02-MAY-2000; 2000JP-0183767.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Kishimoto T;

XX Yano K, Kanazaki K, Inoue Y;

XX WPI; 2001-570288/64.

XX N-PSDB; AAH78275.

XX New gene encoding guanosine triphosphate binding protein coupled

XX receptor, and the protein and antibodies to it, useful for diagnosis

XX and treatment of disease such as brain disease -

XX Claim 2; Page 48-51; 63pp; Japanese.

XX The present sequence represents a human guanosine triphosphate binding

XX protein coupled receptor. The guanosine triphosphate binding

XX protein coupled receptor protein is useful in the diagnosis, prediction

XX and treatment of disease associated with disorders of G protein coupled

XX receptor protein, and may be useful in brain disease and cancers.

XX Sequence 395 AA;

Query Match Best Local Similarity 61.5%; Score 2033; DB 22; Length 395;

Matches 393; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 239 LLSVLCNGLVLLTVFAGGPAFLPPVKFVGAAGANTLTGISCGLASVDALTFGGFSEY 298

DB 1 LLSVLCNGLVLLTVFAGGPAFLPPVKFVGAAGANTLTGISCGLASVDALTFGGFSEY 60

QY 299 GARWETGLGCRATGFLAVLGEASVLLTLAAVQCSVSVSCVRANGKSPSLGSRVAVLG 358

DB 61 GARWETGLGCRATGFLAVLGEASVLLTLAAVQCSVSVSCVRANGKSPSLGSRVAVLG 120

QY 359 CLALAGLAALPLASVEYASPLCLPYAPPEGPAALGFTVALVWMSFCLVVAAGAT 418

DB 121 CLALAGLAALPLASVEYASPLCLPYAPPEGPAALGFTVALVWMSFCLVVAAGAT 180

QY 419 KLYCDLPRGDFEAVWDCAMVRHVAWLTFADGLLYCPVAFSLFASMLGLFPVTPPEAVKSVL 478

DB 181 KLYCDLPRGDFEAVWDCAMVRHVAWLTFADGLLYCPVAFSLFASMLGLFPVTPPEAVKSVL 240

QY 479 LVVLPPLACINPLLYLLENFHRDDLRLRPRAAGDSGPLAAAGELKSSCDSTQALVA 538

DB 241 LVVLPPLACINPLLYLLENFHRDDLRLRPRAAGDSGPLAAAGELKSSCDSTQALVA 300

QY 539 FSDVDLILEASEAGRPGLLETYGFPSTVLLISCOOPGAPRLGSHCVPEEGNHFGNPPSM 598

DB 301 FSDVDLILEASEAGRPGLLETYGFPSTVLLISCOOPGAPRLGSHCVPEEGNHFGNPPSM 360

QY 599 DGEILLRABGSTPAGGGLSGGGGFPSPGLAFASHV 633

DB 361 DGEILLRABGSTPAGGGLSGGGGFPSPGLAFASHV 395

RESULT 11

ID AAW93889 standard; Protein; 907 AA.

XX AAW93889;

DT 25-JUN-1999 (first entry)
 XX Human HG38 protein.
 DE
 XX HG38: human; G-protein coupled glycoprotein hormone receptor; brain;
 KW endocrine system; skeletal muscle; spinal cord; placenta; development;
 KM receptor activity modulator.
 XX
 OS Homo sapiens.
 PN W09915660-A1.
 XX
 PD 01-APR-1999.
 XX
 PF 24-SEP-1998; 98WO-US19979.
 XX
 PR 24-SEP-1997; 97US-0059863.
 XX
 PA (MERT) MERCK & CO INC.
 XX
 PI Bailey WJ, Liu Q, McDonald TP;
 DR WPI: 1999-254711/21.
 DR N-PSDB; AAX23980.
 XX
 XX Human G-protein coupled glycoprotein hormone receptor HG38
 PS Claim 1a; Fig 2; 74pp; English.
 XX
 CC This invention describes a novel human G-protein coupled glycoprotein
 CC hormone receptor, HG38. Glycoprotein hormone receptors are important in
 CC the endocrine system and HG38 may be involved in development and function
 CC of the skeletal muscle, spinal cord, placenta and to a lesser extent, the
 CC brain. The transgenic animal may be useful for studying tissue and
 CC temporal specific expression or activity of the HG38 receptor, as well as
 CC for studying the ability of a variety of compounds to act as modulators
 CC of HG38 receptor activity.
 CC
 SQ Sequence 907 AA;
 Query Match 38.6%; Score 1275; DB 20; Length 907;
 Best Local Similarity 46.9%; Pred. No. 2.5e-107;
 Matches 250; Conservative 96; Mismatches 157; Indels 30; Gaps 8;
 QY 34 KLEDTFQSSLSQALDLSWMAIRSHPEASTLSIKLDTNOLITLPLAGLGLMH 113
 DB 388 EIKVDFFQOLLISRLSLNLAHNAFSLTSLIKLIDSSNLSFPTGLHGLTH 447
 QY 114 LKLGKMLALSOAFSKDSFKRLILEVYAYQCCPYGMCASFRAQGM-----EEDLH 167
 DB 448 LKLTGNHALOSLISSENPFLKVIEMPRAYQCCAFGVCENAYKISQMNKGDNSMDLH 507
 QY 168 LDDESSKRLPLGLARQAEHYNODLDELQLEM-EDSKRPHSVQCSPTGPKPCYLEE 226
 DB 508 -----KKRAGMFQADE---RDLEDFLLDEBDLKLHVSQCSPPGPKRCHLLD 556
 QY 227 SWGIRLAWAIVLISVLCNGIVLTVFAGGAPAPLPVKYVAGALAGANTLTGSCGLAS 286
 DB 557 GWLIRIGWITAVLALCNMLVETVPR-SPLYISPIKLIGVIAAVNMLTGVSSAVLAG 615
 QY 287 VDALTFGQFSEYKARWETGLGCRATGFLAVGSPASVLLTLTAAVQCSVSCVRAVGRS 346
 DB 616 VDAFTFGSFARHGAWMEVGVCHVIGELSTPASSSVFLTLALDEGFSVKYSARETEK 675
 QY 347 PSLGSRVAGVIGCLALAGLALPLASVGEYASPLCLPYAPREGQPAALGFYALVYMN 406
 DB 676 APFSLKAVIILLCALTLALTMVAVPLLGSKYGASPLCLPL--PGEPSITMGYMAVAILLN 733
 QY 407 SFCPLVAVAGATIKLYCDLPRGDEAVVWCAMVRHVAWILIPADGLLYCPVAFSPASMLGL 466
 DB 734 SICPLMTAITAKLYCNLDKGLDENIMDCSMVKHTALTLFTNCILNCYAFALSFSLNL 793
 QY 467 FPVTPFAVSVLVVLPPLPACINPLIYLDFNPHFRDRLRLRPRA-----GDSGLAVAA 521

DB 794 TFIPEVIFILLVWPLPACINPLIYLDFNPHFKEDLVSLRKQYVWTRSKHPSLMSIN 853
 QY 522 AGELEKSCDSTQALVAFSDVDLILEASRGRP---PGLFTYGFPSVTLISC 570
 DB 854 SDDVEKQSCDSTQALVTFNSSITVDLPSSVSPAPYPTESCHLSSVAFVC 906
 RESULT 12
 ID AAY90682
 AC AAY90682 standard; Protein: 907 AA.
 XX
 AC AAY90682;
 XX
 DT 21-AUG-2000 (first entry)
 XX
 DE Human G protein-coupled receptor HG38.
 XX
 DE G protein-coupled receptor; GPCR; constitutively active;
 KW intracellular loop 3; Transmembrane domain 6; drug screening;
 KW agonist; antagonist.
 XX
 OS Homo sapiens.
 PN W0200022129-A1.
 XX
 PD 20-APR-2000.
 XX
 PF 12-OCT-1999; 99WO-US23938.
 XX
 PR 13-OCT-1998; 98US-0170496.
 XX
 PA (AREN-) ARENA PHARM INC.
 XX
 PI Behan DP, Chalmers DT, Liaw CW;
 DR WPI: 2000-329165/28.
 DR N-PSDB; AAA30770.
 XX
 XX Non-endogenous constitutively activated human G protein-coupled
 PT receptors, useful for identifying agonists for use as pharmaceutical
 PT agents -
 PS
 PS Example 1; Page 317-320; 341pp; English.
 XX
 CC The invention relates to constitutively active, non-endogenous versions
 CC of endogenous human orphan G protein-coupled receptors (GPCRs, AAY90643-
 CC AAY90677 and AAY90683-Y90687), and to DNA encoding them (AAA30709-A30743
 CC and AAA30775-A30779). The mutant proteins of the invention contain a
 CC mutation in a portion of the protein comprising intracellular loop 3
 CC (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, X,
 CC is substituted for an endogenous residue in IC3 at a position 16 amino
 CC acids N-terminal of an endogenous proline in TM6 to form a sequence
 CC X-(AA)15-Pro. The endogenous amino acid is selected from Lys, His, Arg
 CC or Ala, and is preferably Lys. When the endogenous residue at this
 CC position is Lys, this residue is replaced by His, Arg or preferably Ala.
 CC The 15 amino acid stretch between the substituted amino acid and the Pro
 CC may be endogenous, non-endogenous, or a mixture of endogenous and
 CC non-endogenous residues. The constitutively active GPCRs are useful for
 CC identifying antagonists, agonists and partial agonists for use as
 CC pharmaceutical agents. The mutant proteins are also useful in research
 CC settings for elucidating the roles of the receptors in normal and
 CC diseased conditions. Antagonists for a particular GPCR are useful for
 CC treating diseases and disorders associated with that receptor. Because
 CC the novel mutant GPCRs are constitutively active, they can be used
 CC directly for screening of compounds without the need for endogenous
 CC ligands. The present sequence represents a human wild-type GPCR referred
 CC to in an exemplification of the invention.
 CC
 SQ Sequence 907 AA;
 Query Match 38.6%; Score 1275; DB 21; Length 907;
 Best Local Similarity 46.9%; Pred. No. 2.5e-107;

Matches 250; Conservative 96; Mismatches 157; Indels 30; Gaps 8;

```

OY 54 KLEXTFESQSLQALDLSWMAIRSHPEAFSTHSLVKLDLTDNQLTTLPLAGLGLMH 113
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 388 EIKVDTEQQLLSLRSLNIAWKIAIHPNAFSTLPSLTKLDLSNLSLSPFTTGLHGLTH 447
OY 114 LKLGKLNALSGAKSDSPFKRLILEVPAYQCCPYGMKASFKAAGW-----EAEELH 167
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 448 LKLGKLNALSGAKSDSPFKRLILEVPAYQCCPYGMKASFKAAGW-----EAEELH 167
OY 168 LDDESSKRPGLGLARQAEHNYDDLEOLEM-EDSKPHPSVOCSPPTGPFKCEYLFE 226
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 508 -----KKDAGMFOADE-----RDLEDFLLDEEDLKAHLSVOCSPSPGPFKEHLDD 556
OY 227 SWGIRLAWAIVLVSLVNCGLVLTVEFAGGAPLPYKFEVVGALAGNTLTGISCGLAS 286
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 557 GMLIRIGWMTIAVLTALCNALVTSTVER-SPLYISPIKLIGVIAAVNMLTGVSASVAVLAG 615
OY 287 VDALTFGQFSEYGARWETGLGCRATGFLAVLGSEASVLLTLTAAVOCVSVCVRAVYKGS 346
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 616 VDAFTFGSFARHGAMWENGCHVIGFLSIFASESSVFLLTLAALRGFSVKYSAKFEYK 675
OY 347 PSLGVRAGVIGCLALAGLAALPLASVEGYGASPLCPYAPPEGAPALGFTVALVMN 406
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 676 APFSSILKVIILLCALLALTMMAVPLLGSKYKASPLCLPL--PFGEPSTMGVWVALITLN 733
OY 407 SFCELVVAGAYIKLYCDLPKRGFEAVWDCAMVRHVAMLIFADGLLYCPVAFSFAAMLGL 466
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 734 SLCEFLMTIATYTKLYCNLDKGDLENIMDCSMVKHIALFTNCILNCVAFISFESSLINL 793
OY 467 FPVTPRAVKSVLVLPPLPACLNPLLYLLENPHKEDLVLSRKQYVWTRSKHPSLMSIN 521
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 794 TFIISPEVVKFILLVVPPLPACLNPLLYLLENPHKEDLVLSRKQYVWTRSKHPSLMSIN 521
OY 522 AGELEKSSCDSTQALVAFSDVDLLEASVGRP-----PGLFTYGPSPVTLISC 570
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 854 SDOVEKQSCDSTQALVAFSDVDLLEASVGRP-----PGLFTYGPSPVTLISC 570

```

RESULT 13

AA42169
ID AAY42169 standard; Protein: 693 AA.

```

XX AC AAY42169;
XX 20-DEC-1999 (first entry)
XX DE Human LGR5 protein sequence.
XX KW Human; LGR4; LGR5; LGR7; G-protein coupled receptor; gene therapy;
XX KM extracellular leucine rich repeat region; mapping; identification.
XX OS Homo sapiens.
XX PN WO9948921-A1.
XX PD 30-SEP-1999.
XX PF 25-MAR-1999; 99WO-US06573.
XX PR 26-MAR-1998; 98US-0079501.
XX PA (STRD ) UNIV LELAND STANFORD JUNIOR.
XX PA (ORGA ) ORGANON NV.
XX PI Hsueh AJW, Hsu SY, Liang S, Van Der Spek PJ;
XX DR WPI: 1999-591074/50.
XX DR N-PSDB; AA25344.
XX PT New G-protein coupled receptors, useful for identifying their own
XX ligands -

```

PS Claim 2; Fig 2; 54pp; English.

XX The present sequence represents the human G-protein coupled receptor
CC having extracellular leucine rich repeat regions, designated LGR5.
CC The LGR4, LGR5 and LGR7 proteins are used to identify ligands for
CC the receptor. The polypeptides and/or polynucleotides are also useful
CC for homologous or related genes, producing compositions that modulate
CC the expression or function of the receptors, gene therapy, mapping
CC functional regions of the receptors, studying associated physiological
CC pathways, in vivo prophylactic and therapeutic purposes, as immunogens
CC for producing antibodies, and for identifying biologically active
CC agents. The polypeptides contain a G-protein coupled seven
CC transmembrane region and a leucine rich repeat extracellular domain.
CC These regions capture and facilitate optimal orientation of its ligand.
CC The proteins are also expressed in diverse tissues.

SO Sequence 693 AA;

Query Match 38.5%; Score 1273; DB 20; Length 693;
Best Local Similarity 46.9%; Pred. No. 2.6e-107;

Matches 250; Conservative 96; Mismatches 157; Indels 30; Gaps 8;

```

OY 54 KLEXTFESQSLQALDLSWMAIRSHPEAFSTHSLVKLDLTDNQLTTLPLAGLGLMH 113
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 174 EIKVDTEQQLLSLRSLNIAWKIAIHPNAFSTLPSLTKLDLSNLSLSPFTTGLHGLTH 233
OY 114 LKLGKLNALSGAKSDSPFKRLILEVPAYQCCPYGMKASFKAAGW-----EAEELH 167
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 234 LKLGKLNALSGAKSDSPFKRLILEVPAYQCCPYGMKASFKAAGW-----EAEELH 167
OY 168 LDDESSKRPGLGLARQAEHNYDDLEOLEM-EDSKPHPSVOCSPPTGPFKCEYLFE 226
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 294 -----KKDAGMFOADE-----RDLEDFLLDEEDLKAHLSVOCSPSPGPFKEHLDD 342
OY 227 SWGIRLAWAIVLVSLVNCGLVLTVEFAGGAPLPYKFEVVGALAGNTLTGISCGLAS 286
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 343 GMLIRIGWMTIAVLTALCNALVTSTVER-SPLYISPIKLIGVIAAVNMLTGVSASVAVLAG 401
OY 287 VDALTFGQFSEYGARWETGLGCRATGFLAVLGSEASVLLTLTAAVOCVSVCVRAVYKGS 346
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 402 VDAFTFGSFARHGAMWENGCHVIGFLSIFASESSVFLLTLAALRGFSVKYSAKFEYK 461
OY 347 PSLGVRAGVIGCLALAGLAALPLASVEGYGASPLCPYAPPEGAPALGFTVALVMN 406
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 462 APFSSILKVIILLCALLALTMMAVPLLGSKYKASPLCLPL--PFGEPSTMGVWVALITLN 519
OY 407 SFCELVVAGAYIKLYCDLPKRGFEAVWDCAMVRHVAMLIFADGLLYCPVAFSFAAMLGL 466
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 520 SLCEFLMTIATYTKLYCNLDKGDLENIMDCSMVKHIALFTNCILNCVAFISFESSLINL 579
OY 467 FPVTPRAVKSVLVLPPLPACLNPLLYLLENPHKEDLVLSRKQYVWTRSKHPSLMSIN 521
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 580 TFIISPEVVKFILLVVPPLPACLNPLLYLLENPHKEDLVLSRKQYVWTRSKHPSLMSIN 521
OY 522 AGELEKSSCDSTQALVAFSDVDLLEASVGRP-----PGLFTYGPSPVTLISC 570
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 640 SDOVEKQSCDSTQALVAFSDVDLLEASVGRP-----PGLFTYGPSPVTLISC 570

```

RESULT 14

AA90687
ID AAY90687 standard; Protein: 907 AA.

```

XX AC AAY90687;
XX 21-AUG-2000 (first entry)
XX DE Human mutant G protein-coupled receptor HG38 (V765K).
XX KW G protein-coupled receptor; GPCR; constitutively active;
XX KM intracellular loop 3; transmembrane domain 6; drug screening;
XX agonist; antagonist; mutant; mutelin.

```


OS Homo sapiens.
 OS Synthetic.
 PN WO200022129-A1.
 PD 20-APR-2000.
 XX 12-OCT-1999; 99WO-US23938.
 XX 13-OCT-1998; 98US-0170496.
 XX (AREN-) ARENA PHARM INC.
 PA Behan DP, Chalmers DT, Liaw CW;
 PI WPI: 2000-329165/28.
 DR N-PSDB; AAA30779.
 XX Non-endogenous constitutively activated human G protein-coupled
 PT receptors, useful for identifying agonists for use as pharmaceutical
 PT agents -
 PS Example 2; Page 332-335; 341pp; English.
 XX The invention relates to constitutively active, non-endogenous versions
 CC of endogenous human orphan G protein-coupled receptors (GPCRs, AAY90643-
 CC AAY90677 and AAY90683-Y90687), and to DNA encoding them (AAA30709-A30743
 CC and AAA30775-A30779). The mutant proteins of the invention contain a
 CC mutation in a portion of the protein comprising the intracellular loop 3
 CC (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, X,
 CC is substituted for an endogenous residue in IC3 at a position 16 amino
 CC acids N-terminal of an endogenous proline in TM6 to form a sequence
 CC X-(AA)15-Pro. The endogenous amino acid is selected from Lys, His, Arg
 CC or Ala, and is preferably Lys. When the endogenous residue at this
 CC position is Lys, this residue is replaced by His, Arg or preferably Ala.
 CC The 15 amino acid stretch between the substituted amino acid and the Pro
 CC may be endogenous, non-endogenous, or a mixture of endogenous and
 CC non-endogenous residues. The constitutively active GPCRs are useful for
 CC identifying antagonists, agonists and partial agonists for use as
 CC pharmaceutical agents. The mutant proteins are also useful in research
 CC settings for elucidating the roles of the receptors in normal and
 CC diseased conditions. Antagonists for a particular GPCR are useful for
 CC treating diseases and disorders associated with that receptor. Because
 CC the novel mutant GPCRs are constitutively active, they can be used
 CC directly for screening of compounds without the need for endogenous
 CC ligands. Sequences AAY90643- AAY90677 and AAY90683-Y90687 the mutant
 CC human GPCRs of the invention.
 CC
 SQ Sequence 907 AA;
 XX
 Query Match 38.4%; Score 1269; DB 21; Length 907;
 Best local Similarity 46.7%; Pred. No. 8.9e-107;
 Matches 249; Conservative 96; Mismatches 158; Indels 30; Gaps 8;
 QY 54 KLEKDTFQSLSSLDALDLSNNAITRPEAFSTLHSLVKLDTJNOLTTPLAGIGGLMH 113
 DB EIKVDTFQSLSSLDALDLSNNAITRPEAFSTLHSLVKLDTJNOLTTPLAGIGGLMH 447
 QY 114 LKLGNNALSGAFKDSPPKRLILEVPAAYCCPCPYGMCASFKAAGOW-----EADDLH 167
 DB 448 LKLGNNALSGAFKDSPPKRLILEVPAAYCCPCPYGMCASFKAAGOW-----EADDLH 507
 QY 168 LDDESSKRPGLARQAEHHYDDELQLEM-EDSKPHPSVQCSPPGPKCEYLF 226
 DB 508 -----KKDAGMFQAOE-----RDLEDPLDFEEDLKALHSVQCSPPGPKCEYLF 556
 QY 227 SWGIRLAWMAVLVSLVLCNGVLTVEFGAGAPLPPEVKEVVGALGANTLIGISGLLAS 286
 DB 557 GMLIRIGWTIAVLATLTCNALVTSTVER-SPLYYSIKLLIGVIAVAMLTGSSAVIAG 615
 QY 287 VDAITFGGFSEYGAEMETGLCRATGFLAVLGSEASVLLTLTAAVQCSVSVCAVAYGKS 346
 DB 616 VDAITFGGFSEYGAEMETGLCRATGFLAVLGSEASVLLTLTAAVQCSVSVCAVAYGKS 675

QY 347 PSLSGVRAGVIGCLALAGIAALPLASVGEYASPLCLPYAPPEGQAPALGFTVALVMNN 406
 DB 676 APFSSLKVTILLICALLALMAAVPLIGSKYASPLCLPL--PGEPSYMGVMAVLITLIN 733
 QY 407 SFCFLVAGATIKLYCDLPRGDFEAVWDCAMVRHVAWILFADGLLYCPAFYAFSFMGL 466
 DB 734 SLCEFLMTIATYKLYCNLDKGDLENIMDCSMKKHIALTLFTCNILNCPAFYAFSFMGL 793
 QY 467 FVTPPEAVKSVLLVLPPLACNPLLYLLFNPHFRDRLRLRPA-----GSGPLATA 521
 DB 794 TPISPEVIKFILLVVPPLACNPLLYLLFNPHFKEDLVSLRKQYVWTRSKHPSLMSIN 853
 QY 522 AGELEKSSCDSTQALVAEFDVLLLEASAGRP-----PLETYGPPSYVLIS 570
 DB 854 SDVEKQSCDSTQALVTFSSITTYDLPSSVPSPAYPTESCHLSSVAFPVC 906
 RESULT 15
 AAB95066
 ID AAB95066 standard; Protein; 230 AA.
 AC AAB95066;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:16921.
 XX
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 OS Homo sapiens.
 XX
 XX EPI074617-A2.
 PN 07-FEB-2001.
 PD 28-JUL-2000; 2000EP-0116126.
 PF 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELT-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI: 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 PS Claim 8; SEQ ID 16921; 2537p + CD ROM; English.
 XX
 The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a 3'-end sequence, where the
 CC polynucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX
SQ Sequence 230 AA;

Query Match 36.5%; Score 1206; DB 22; Length 230;
Best Local Similarity 99.6%; Pred. No. 7.3e-102;

Matches 229; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 404 MNNSFCFLVYAGAYIKLYCDLPRGDPEAVWDCAMVRHVAWLIFADGLYCPYAFISFASM 463
|||
Db 1 MNNSFCFLVYAGAYIKLYCDLPRGDPEAVWDCAMVRHVAWLIFADGLYCPYAFISFASM 60
|||
QY 464 LGLEPYTPPEAVKSVLLVLLPACLNPLLYLLFNPHERDLRLRPRAGDSGLAYAAAG 523
|||
Db 61 LGLEPYTPPEAVKSVLLVLLPACLNPLLYLLFNPHERDLRLRPRAGDSGLAYAAAG 120
|||
QY 524 ELEKSSCDSTQALVAFSDVDLILEASEAGRPGLETYGFPSVTILISCOOPGAPRLGSHC 583
|||
Db 121 ELEKSSCDSTQALVAFSDVDLILEASEAGRPGLETYGFPSVTILISCOOPGAPRLGSHC 180
|||
QY 584 VEPGHNHFGNPPQPMDEGLLLRAEGSTPAGGGLSGGGGFQPSGLAFASHV 633
|||
Db 181 VEPGHNHFGNPPQPMDEGLLLRAEGSTPAGGGLSGGGGFQPSGLAFASHV 230
|||

Search completed: November 8, 2002, 19:34:26
Job time : 37.8291 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2002, 19:32:50 ; Search time 36.5581 Seconds
(without alignments)
4148.213 Million cell updates/sec

Title: US-09-851-595-8
Perfect score: 3850
Sequence: 1 GLHNLETTDLNKNKLEFPV.....GGLSGGGGFPQSLAFASHV 736

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 20604715 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.ricent:*
12: sp.virus:*
13: sp.vertibrate:*
14: sp.unclassified:*
15: sp.rvivirus:*
16: sp.bacteriophage:*
17: sp.archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3850	100.0	928	4 Q9BYD7	Q9BYD7 homo sapien
2	2087	54.2	549	11 Q8R301	Q8R301 mus muscula
3	840.5	21.8	1360	5 Q9ND11	Q9ND11 drosophila
4	839.5	21.8	1050	5 Q9BN18	Q9BN18 drosophila
5	827.5	21.5	1300	5 Q9NKM6	Q9NKM6 drosophila
6	766.5	19.9	1012	5 Q95Y16	Q95Y16 asterina pe
7	766.5	19.9	1280	5 Q95Y17	Q95Y17 asterina pe
8	545.5	14.2	701	13 Q9DGC6	Q9DGC6 oreochromis
9	504.5	13.1	688	11 Q64183	Q64183 rattus sp.
10	490.5	12.7	695	11 Q8R428	Q8R428 capra porce
11	487	12.6	724	13 Q9PVP0	Q9PVP0 oncorhynch
12	482.5	12.5	658	13 Q9PVP9	Q9PVP9 oncorhynch
13	478.5	12.4	779	13 Q918N7	Q918N7 morone saxa
14	475.5	12.4	693	13 Q9DGC5	Q9DGC5 oreochromis
15	475.5	12.4	829	5 Q9VEG4	Q9VEG4 drosophila
16	475.5	12.4	831	5 Q94979	Q94979 drosophila

17	470	12.2	778	13 Q98TF4	Q98TF4 oreochromis
18	468.5	12.2	831	5 Q8SX01	Q8SX01 drosophila
19	460.5	12.0	763	6 Q9BGN4	Q9BGN4 felis silve
20	459	11.9	696	13 Q9DGF5	Q9DGF5 cynops pyr
21	454.5	11.8	764	11 Q9D697	Q9D697 mus musculu
22	450	11.7	701	6 Q8SPS8	Q8SPS8 bos taurus
23	442.5	11.5	793	13 Q91948	Q91948 oncorhynch
24	436.5	11.3	764	6 Q8SP56	Q8SP56 sus scrofa
25	436.5	11.3	764	6 Q8SP59	Q8SP59 sus scrofa
26	434	11.3	662	13 Q9PW16	Q9PW16 ciarias gar
27	433	11.2	739	6 Q9B555	Q9B555 sus scrofa
28	432.5	11.2	601	13 Q42500	Q42500 meleagris g
29	431	11.2	929	5 Q18759	Q18759 caenorhabdi
30	428.5	11.1	814	13 Q91949	Q91949 oncorhynch
31	425	11.0	662	13 Q98T84	Q98T84 ictalurus p
32	420.5	10.9	696	13 Q98T85	Q98T85 ictalurus p
33	420	10.9	699	4 Q15966	Q15966 homo sapien
34	413	10.7	701	4 Q14751	Q14751 homo sapien
35	382	9.9	533	4 Q9P1V4	Q9P1V4 homo sapien
36	371.5	9.6	470	4 Q43200	Q43200 homo sapien
37	366.5	9.5	410	4 Q16225	Q16225 homo sapien
38	338	8.8	673	13 Q90WP8	Q90WP8 podarcis si
39	328	8.5	590	6 Q27986	Q27986 bos taurus
40	311	8.1	293	6 Q95ME7	Q95ME7 canis fam11
41	284	7.4	335	6 Q46387	Q46387 mustela vis
42	270.5	7.0	307	6 Q95159	Q95159 alluropoda
43	259.5	6.7	228	11 Q9BSJ8	Q9BSJ8 rattus norv
44	255.5	6.6	307	13 Q90XC7	Q90XC7 salmo salar
45	255.5	6.6	332	13 Q8QFN6	Q8QFN6 elaphe quad

ALIGNMENTS

RESULT 1
Q9BYD7 PRELIMINARY; PRT; 928 AA.

AC Q9BYD7;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE VTS20631 (Fragment).
GN VTS20631.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Okaze H., Hayashi A., Kozuma S., Saito T.;
RT "A member of g-protein coupled receptor family."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB049405; BAB39854.1; -
DR HSSP; P23945; IXON.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR003591; LRR_ryp.
DR Pfam; PF00560; LRR_13.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PRINTS; PR00019; LRR_RICHPT.
DR SMART; SM00370; LRR_9.
DR SMART; SM00369; LRR_TYP; 14.
FT NON_TER
SQ SEQUENCE 928 AA; 100487 MW; 4C3364ADEA89C463 CRC64;

Query Match 100.0%; Score 3850; DB 4; Length 928;
Best Local Similarity 100.0%; Pred. No. 2.5e-292;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLHNLETTDLNKNKLEFPVALRTIGRLQELGFHNNNKAIPKAKMGAPLLQTHFYFN 60
DB 193 GLHNLETTDLNKNKLEFPVALRTIGRLQELGFHNNNKAIPKAKMGAPLLQTHFYFN 252

```

QY 61 PLOFVGRSAFOYLPKLTHTLSNGAMDIQEPDLKGTSTLETLTLTRAGTRILPSCMCOL 120
DB 253 PLOFVGRSAFOYLPKLTHTLSNGAMDIQEPDLKGTSTLETLTLTRAGTRILPSCMCOL 312
QY 121 PRLRVLELSHQBELPFLSHRCQKLEETIGLOHNRITWEGADTFQSLSLQALDLSSMNAIR 180
DB 313 PRLRVLELSHQBELPFLSHRCQKLEETIGLOHNRITWEGADTFQSLSLQALDLSSMNAIR 372
QY 181 SIHPEAFSTHSIVKDLDTLONOLTLPLAGIGLMHLTKLNALSCAFSKDSFPKRLT 240
DB 373 SIHPEAFSTHSIVKDLDTLONOLTLPLAGIGLMHLTKLNALSCAFSKDSFPKRLT 432
QY 241 EYVYAYQCCPCYGCASFKAASGQWEADLHLDDDESSKRPGLGLARQAEHYDDDLQ 300
DB 433 EYVYAYQCCPCYGCASFKAASGQWEADLHLDDDESSKRPGLGLARQAEHYDDDLQ 492
QY 301 LEMEDSKPHSVQCCSPPEGFPCPEYLFESWGRFLAVMAIYVLSVLCNGVLLTVFAGCP 360
DB 493 LEMEDSKPHSVQCCSPPEGFPCPEYLFESWGRFLAVMAIYVLSVLCNGVLLTVFAGCP 552
QY 361 APLPVKFEVYVAGTAGANTLTGISCGLASVDALTFQGFSEYGARWETGLGRATGFLAVL 420
DB 553 APLPVKFEVYVAGTAGANTLTGISCGLASVDALTFQGFSEYGARWETGLGRATGFLAVL 612
QY 421 GSEASVLLTLTAAVQCSVSVSCYRAVKGSPSLGSRAGVLGCLALAGLAAALPLASVGEY 480
DB 613 GSEASVLLTLTAAVQCSVSVSCYRAVKGSPSLGSRAGVLGCLALAGLAAALPLASVGEY 672
QY 481 GASPLCLPAPPEGQPAALGFTVALVMMNSFCFLVYAGATIKLYCDLPKDDFEVWVCCAM 540
DB 673 GASPLCLPAPPEGQPAALGFTVALVMMNSFCFLVYAGATIKLYCDLPKDDFEVWVCCAM 732
QY 541 VHVAVMLIFADDLTLCVPAFLSPASMLGLFPVTPPEAVKSVLVVLEPLPACLNPLLYLFFN 600
DB 733 VHVAVMLIFADDLTLCVPAFLSPASMLGLFPVTPPEAVKSVLVVLEPLPACLNPLLYLFFN 792
QY 601 PHFRDRLRLRPRAGDSGFLAYAAAGELKSSQDSQALVAFSDVLLLEASAGRPGL 660
DB 793 PHFRDRLRLRPRAGDSGFLAYAAAGELKSSQDSQALVAFSDVLLLEASAGRPGL 852
QY 661 ETYGPPSVTLISCOQPGAPRLGSHCVPEEGNHFGNPPQPSMDGELLRAEGSTPAGGGLS 720
DB 853 ETYGPPSVTLISCOQPGAPRLGSHCVPEEGNHFGNPPQPSMDGELLRAEGSTPAGGGLS 912
QY 721 GGGGPPSGGLAFASHV 736
DB 913 GGGGPPSGGLAFASHV 928

```

RESULT 2

```

Q8R301 PRELIMINARY: PRT: 459 AA.
AC Q8R301;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DE Hypothetical 47.9 kDa protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC026896; AAH26896.1; -.
KW Hypothetical protein.
FT NON-TER
SQ SEQUENCE 459 AA: 47889 MW: F0100BF073E81762 CRC64;

```

Query Match 54.2%; Score 2087; DB 11; Length 459;
Best Local Similarity 87.68; Freq. No. 7, 5e-155;

Matches 402; Conservative 20; Mismatches 37; Indels 0; Gaps 0;

```

QY 278 KRLGLGLARQAEHYDDDLQEMEDSKPHSVQCCSPPEGFPCPEYLFESWGRFLAV 337
DB 1 KRLGLGLARQAEHYDDDLQEMEDSKPHSVQCCSPPEGFPCPEYLFESWGRFLAV 60
QY 338 WALVLSVLCNGVLLTVFAGPAPLPVYKFEVYVAGTAGANTLTGISCGLASVDALTFQ 397
DB 61 WALVLSVLCNGVLLTVFAGPAPLPVYKFEVYVAGTAGANTLTGISCGLASVDALTFQ 120
QY 398 FSEYGARWETGLGRATGFLAVMAIYVLSVLCNGVLLTVFAGCP 457
DB 121 FAEYGARWESGLGCAQTGFLAVGDSAVLLTLAVQCSIYVTCVARYKAPSPGSVRA 180
QY 458 GVLGCLALAGLAAALPLASVGEYASPLCLPAPPEGQPAALGFTVALVMMNSFCFLVVA 517
DB 181 GALGCLALAGLAAALPLASVGEYASPLCLPAPPEGQPAALGFTVALVMMNSFCFLVVA 240
QY 518 GAYIKLYCDLPKDDFEVWVCCAMVHVAVMLIFADGLLYCPVAFLSFASMLGLFPVPEAV 577
DB 241 GAYIKLYCDLPKDDFEVWVCCAMVHVAVMLIFADGLLYCPVAFLSFASMLGLFPVPEAV 300
QY 578 KSVLLVPLPACLNPLLYLFFNPHFRDRLRLRPRAGDSGFLAYAAAGELKSSQDSQ 637
DB 301 KSVLLVPLPACLNPLLYLFFNPHFRDRLRLRPRAGDSGFLAYAAAGELKSSQDSQ 360
QY 638 ALVAFSDVLLLEASAGRPGLGTYGPPSVTLISCOQPGAPRLGSHCVPEEGNHFGN 697
DB 361 ALVAFSDVLLLEASAGRPGLGTYGPPSVTLISCOQPGAPRLGSHCVPEEGNHFGN 736
QY 698 QPSMDGELLRAEGSTPAGGGLGFGPPSGGLAFASHV 736
DB 421 QPPMKGELLRAEGSTPAGGGLGFGPPSGGLAFASHV 459

```

RESULT 3

```

Q9NDI1 PRELIMINARY: PRT: 1360 AA.
AC Q9NDI1;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE Glycoprotein hormone receptor II.
GN RK OR CG8930.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CANTON S; TISSUE-WHOLE ANIMAL;
RX MEDLINE=20359836; PubMed=10899142;
RA Eriksen K.K., Hauefer F., Schiott M., Pedersen K.-M., Soendergaard L.,
RA Grimelinkhuijzen C.J.P.;
RT "Molecular Cloning, Genomic Organization, Developmental Regulation,
RT and a Knock-Out Mutant of a Novel Leu-Rich Repeats-containing G
RT Protein-Coupled Receptor (DLCR-2) from Drosophila melanogaster.";
RL Genome Res. 10:924-938(2000).
DR EMBL; AF142343; AAFF6608.1; -.
DR HSSP; Q57815; ID3Y.
DR Flybase; FBgn0003255; rK.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 14.
DR PRINTS; PR00237; GPCR_RHODPSN.
DR SMART; SM00370; LRR; 2.
DR SMART; SM00369; LRR_TYP; 5.
DR PROSITE; PS50262; G_PROTEIN_RECPT_F1_2; 1.
KW Receptor.

```


DB 784 AKFTYVFLPLNSCCNDFLYAINTKQFKKDCVFLCKHFEESRYVGGGPGRCGVAATTKR 843
 QY 631 SSCDSTQALVAFSVDVLD---ILEASEGRRPGLETTFPSVTLSCQOPAPRL 682
 DB 844 G-----DLPPFLPAAVAHPGCR-----CLRMLPSPMPWHKME 879
 RESULT 5
 O9NKD6 PRELIMINARY; PRT: 1300 AA.
 ID O9NKD6
 AC O9NKD6; O9YU3;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 GN Hypothetical 144.0 kDa protein (Rk gene product).
 OS Rk OR CG8930.
 OC Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RA MEDLINE=99403001; PubMed=10471707;
 RA Ashburner M., Misra S., Roote J., Lewis S.E., Blazej R., Davis T.,
 RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
 RA Hong L., Houston K., Hoskins R., Johnson T., Martin C., Moshrefi A.,
 RA Palazzolo M., Reese M.G., Spradling A., Tang G., Wan K., Whitehead K.,
 RA Gelniker S., Rubin G.M.;
 RA "An exploration of the sequence of a 2.9-Mb region of the genome of
 RT Drosophila melanogaster: the Adh region.";
 RL Genetics 153:179-219(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RA Gelniker S.E., Aghayani A., Arcaina T.T., Baxter E., Blazej R.G.,
 RA Butenoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
 RA Farian D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
 RA Houston K.A., Hummasti S.R., Karia K., Kearney L., Kim E., Lee B.,
 RA Lewis S., Li P., Lomolan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
 RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeita A.,
 RA Setti H., Snit E., Svirskas R.R., Wan K.H., Weinburg T., Zhang R.,
 RA Zieran L.L., Rubin G.L.;
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RA MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Gelniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.E., Aghayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Bailett R.W., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakoff S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke G., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Kaulush F., Karpis G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Jaisli B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laslo P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeter F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner C., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wattaman D.A., Weinstock G.M., Welschbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Zhong W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003408; AAF4846.1;
 DR EMBL: AE003642; AAF5367.2;
 DR Flybase: FBgn0003255; tk.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR003591; LRR_typ.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR_10.
 DR PRINTS: PR00237; GPCRHOPOSN.
 DR SMART: SM00370; LRR_TYP.
 DR SMART: SM00369; LRR_TYP; 4.
 DR PROSITE: PS50262; G-PROTEIN_RECPT_FL2; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 1300 AA; 144031 MW; BAB9E39F42FA0B3 CRC64;

Query Match 21.5%; Score 827.5; DB 5; Length 1300;
 Best Local Similarity 29.4%; Pred. No. 9,4e-56;
 Matches 224; Conservative 128; Mismatches 296; Indels 115; Gaps 19;

QY 5 LEPTIDLVNKRLOEPPVAIRTLGRLOELGFNNNNIKAIPEKAFKGNPLQTIHYNDPIQ 64
 DB 301 LKYLELDNDNLSSLPBGLSKLSQLELSTSNLRWINDLEPRS--MQLMDRANPLST 358
 QY 65 VGRSAFOYLPKHTLSLNGAMDIOEPPDKGTSTSLTLTRAGIRLPSGMCQOLPR 124
 DB 359 ISGARGSKLKLTLSDVRLRSPEPELACHALILKIDRAGIOEVANLCROPRILK 418
 QY 125 VELSHNOIEELPS--LHNCQKLEEGIQHNRIMEGADTFQSSQLV----- 172
 DB 419 SLDLSSNQIEKIQGKFNQKQINDLLEYSNNIKALPDQAFGIPLQILVYSIIRIHS 478
 QY 173 DLSMAIRSIHPEASTLSLVLKLDITDQTLTLPPLAGLGLMHLKGNLALSOAFSKD 232
 DB 479 DLGWEISTIRKEAFSGFTALDNLGNINIFELPESGIRALLHKTNPKLREFPPD 538
 QY 233 SFPKRLILEVPAYOC--PYGMCASFVKAS----- 261
 DB 539 TFPRIOTLILSAVHCACFLPLVAMSSOKTISOEAVLPSPDAEFDMTLNNSMINTP 598
 QY 262 -----GQRAEDLHLDDESSRRPLGLLRAQENHYDQDLE----- 298
 DB 599 QMHNLKQGLASMDPWETA--INFNEEQLOTQGOIATSMEEFEBHDVSGEATGYCF 657
 QY 299 -----LOLEMDSKRPHSPSCSPRPFPKPEYFEESMGRILAWAIVLVSLCNGVL 353
 DB 658 GTGLFGSMSTEDFQ--GSVQCILPFGFPLCADLFDWTLRCGVVWVLLSLGNGTVE 716
 QY 354 TVPAGGAPLPKPKFVYGAAGANTLISGCLIASVDALTFQOFSYSGARWEGGCGRA 413
 DB 717 VLIC--SRKMDVPRFICNLAAADFFWIGLITAIYDANTLDEFRFAIPWQASVLCQ 775
 QY 414 TGPLAVGSEASVLLTLAAVQCSVSCRAIGKSPSLGSAVAGVGL--ALAGIAA 471
 DB 776 SGPLAVLSLSL--YTLAVITIERNYATLTHLTKRLSLKQAGYIMSGWFAFLMAL 833
 QY 472 LPLASGEVGSAPLCPVAPPEGAPALGFTVALVMMNSFCFLVAVGATIKYLCDDPRD 531
 DB 834 MPLVGSYDRKRAVCLPFEETTG--PASLTYYVLSIMFINCAFLITLMGCTLKMWAL-RG- 890

OY 532 FEAVW---DCAMRVHVMILTFADGLICYPVAFSLFASMLGLFPVTPPEAVKSVLLVPLP 568
 DB 891 -SOANNNTDSRIKAKMALVFTDFLCWSPILAFFSITAFIAGLQSLSDQAKIFVYVPLN 949
 OY 589 ACNLNLLYLFPNFRDDLRLRLPR-----AGDSGLYATAAGLELKSQSDSTQALVAF 642
 DB 950 SCNNPLLYAIMKKOKKCVTLCKHFEESSRVVGGGCGRCAGVARTKRG----- 998
 OY 643 SDVDL---ILFASGAGRPGLFETGFPSTYLSLSCQOCPARPLE 682
 DB 999 ---DLPPPLPAAVAHAPGCR-----CLRMPLSPMPNMHKKME 1033
 RESULT 6
 O95Y16 PRELIMINARY; PRT; 1012 AA.
 AC O95Y16; 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 21, last annotation update)
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
 DE Glycoprotein hormone receptor.
 GN AGPGR.
 OS Asterina pectinifera (Starfish).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
 OC Asteroidea; Valvatacea; Valvatida; Asterinidae; Asterina.
 OC NCBI_TaxID=7594;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mita M., Hirai T., Oba Y., Yoshikuni M., Nagahama Y.;
 RT "cDNA cloning and functional analysis of a novel member of the
 RT glycoprotein hormone receptor family from a starfish Asterina
 RT pectinifera." (MAY-2001) to the EMBL/Genbank/DBJ databases.
 RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.
 RT EMBL: AB061862; BAB68209.1; ASTRNA_11gase11.
 DR InterPro: IPR002106; ASTRNA_11gase11.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Nterm.
 DR Pfam: PF00560; 7tm_1; 1.
 DR Pfam: PF00560; LRR_14.
 DR SMART: SM00013; LRRNT; 1.
 DR PROSITE: PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
 DR PROSITE: PS00237; G_PROTEIN_RECP_FL_1; UNKNOWN_1.
 DR PROSITE: PS50262; G_PROTEIN_RECP_FL_2; 1.
 KW Receptor.
 SO SEQUENCE. 1012 AA; 112623 MW; 52A70E7A88C46B0A CRC64;
 Query Match 19.9%; Score 766.5; DB 5; Length 1012;
 Best Local Similarity 28.8%; Pred. No. 4e-51;
 Matches 197; Conservative 104; Mismatches 287; Indels 95; Gaps 13;
 OY 5 LETTDLNANKLOEPP-VAIRLGLRLOELGFHNNNKAIPKAFKMGN----- 49
 DB 167 LHHNLNDHNLSEVPTALHLSNLRILHLSNIPYVVDHAFNSHLLLEILRNKKT 226
 OY 50 -----PLQTIHFYDNPIDQVGSARFOYLPRKHLTSLNSGAMDIOEPDLKGTSL 99
 DB 227 HLSAHAFAGLPNMLLEFLGNSITSIATFAFRNLPALRLNVILEVKNLSVFPDLTGTSL 286
 OY 100 EITLIRAGIRLPSGMCQDLRLVLELSHNOIELEPILSRCKLE----- 146
 DB 287 EHLGIRCSLRAPNFCNNMTGLTSLNHNMLLEGPSLSKCSLKVHLGTNKTSL 346
 OY 147 -----ELGLOHNIMEIGADTFSLSLQALDLSMNAIRSIHPFASTLSLVK 195
 DB 347 GQPSGLHLDYLDQLLENDISYIPNAPAFSLDITLSNNTIREIDSOAFAPCTSLQY 406
 OY 196 LDTLNDLTLPLAGLGLMLKLNALSOAFSKDPSPKRLILEVPAYOCCPYGMA 255
 DB 407 LDISNNSFVLPYTAGLOMLKIRTYNEQLEDFPPSELPSTEIATAPYHCCEYELA 466
 OY 256 SFFKASQGWAEADLHDEESSKRPGLGLARQAEHNYDODLEQLQEMDS----- 306

DB 467 BEYLKS-----LADRPNISETTYWAGSVDPDYNTMTFIDINSEWIDSJFGGSLI 517
 OY 307 -----KPHSVOCSPPTGPFKPCLELFESWIGIRLVAIAVLISVLCGLV 352
 DB 518 GSPYLSGYSRLVPH-NISCRKPPGPFPCMDLFGSMLRICVMVLFLALIGNAIVI 576
 OY 353 LTVFAGRAPLPRVKEVVGALAGANTLTGICGLASVDALTFQSESEYCARWETGLGR 412
 DB 577 FVITVSHTKMDVP-REFLONLAFADPFLGVYGLAGVDITSLVFRKFGARWOLSNAGR 635
 OY 413 ATGFLAVLGSSESVLLTLAAVQCSVSVCAVAKGSPSLGSVAGVGLCA--LAGLAA 470
 DB 636 LAGFLAVFSEFSFISYTLTLETFEYAIKHALHLERKML--PHAITMCGWIFSVTAA 693
 OY 471 ALPLASGEYGAASPLCLPAAPREGOPALGFTVALVMNSFCFLVAAAYIKLYCDLPRG 530
 DB 694 VLPVNVSHYHRVPCLPF--DVDITVAKVYGSILILNIAFVIMACYSIYAL-QG 750
 OY 531 DEFAVWDC--AMVRHVMILTFADGLICYPVAFSLFASMLGLFPVTPPEAVKSVLLVPL 587
 DB 751 SH--AMNCNDSRVARRMSLLVFTDFACWAPIAFPSLTAAGRLISLQAKVLTIFVPL 808
 OY 588 PACNPLLYLFPNFRDDLRL 610
 DB 809 NSCANPFLYTLTKQFKKCKTI 831
 RESULT 7
 O95Y17 PRELIMINARY; PRT; 1280 AA.
 AC O95Y17; 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 21, last annotation update)
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
 DE Glycoprotein hormone receptor.
 GN AGPGR.
 OS Asterina pectinifera (Starfish).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
 OC Asteroidea; Valvatacea; Valvatida; Asterinidae; Asterina.
 OC NCBI_TaxID=7594;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mita M., Hirai T., Oba Y., Yoshikuni M., Nagahama Y.;
 RT "cDNA cloning and functional analysis of a novel member of the
 RT glycoprotein hormone receptor family from a starfish Asterina
 RT pectinifera." (MAY-2001) to the EMBL/Genbank/DBJ databases.
 RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.
 RT EMBL: AB061861; BAB68208.1; ASTRNA_11gase11.
 DR InterPro: IPR002106; ASTRNA_11gase11.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Nterm.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR_14.
 DR SMART: SM00013; LRRNT; 1.
 DR PROSITE: PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
 DR PROSITE: PS00237; G_PROTEIN_RECP_FL_1; UNKNOWN_1.
 DR PROSITE: PS50262; G_PROTEIN_RECP_FL_2; 1.
 KW Receptor.
 SO SEQUENCE. 1280 AA; 141700 MW; 0AB0ECC0DD880BA CRC64;
 Query Match 19.9%; Score 766.5; DB 5; Length 1280;
 Best Local Similarity 28.8%; Pred. No. 5.4e-51;
 Matches 197; Conservative 104; Mismatches 287; Indels 95; Gaps 13;
 OY 5 LETTDLNANKLOEPP-VAIRLGLRLOELGFHNNNKAIPKAFKMGN----- 49
 DB 167 LHHNLNDHNLSEVPTALHLSNLRILHLSNIPYVVDHAFNSHLLLEILRNKKT 226
 OY 50 -----PLQTIHFYDNPIDQVGSARFOYLPRKHLTSLNSGAMDIOEPDLKGTSL 99
 DB 227 HLSAHAFAGLPNMLLEFLGNSITSIATFAFRNLPALRLNVILEVKNLSVFPDLTGTSL 286

```

QY 100 EITLFRAGIRLPLSGMCOQLPRLVLELSHNOIEELPSLHRCCKE----- 146
DB 287 EHLIGKCSIRAIIPANCMDNTGLTSLMNNLLEGLPSLCKSSKSLHLGINKLISLE 146
QY 147 -----EIGLOHNRIRWEIGADTFQSLSLOALDLSWMNIRSHPEAFSTLASHYK 195
DB 347 GQPSGSHLDYDQLENDISYIPADAFQSLSHLDLTLSTNNITREIDSOAFPCSTLOY 195
QY 196 LDIITDQTLPLAGLGLMHLTKGNLALSOAFKDSFPRKILEVPAVOCCPYGMCA 255
DB 407 LDISSNSFPPLPTAGLOMLKITTYNEQLEDPPPSSELSITEITATVAPYHCCEYIELA 255
QY 256 SFPKASGQMEADLHLDDESSKRLGLLAROENHYDDLDLEQLEMDS----- 306
DB 467 EELYS-----LADRNISSETTYMASGSVPYNNPTFINESMTDSIFGGSISI 306
QY 307 -----KPHSVOCSPPPGPKPCEYLFESMGIRLAVMAIVLLSVLCNGVL 352
DB 518 GSPYLSGNYSLVLPN-NISCRPKPFPMPGMDLFGSMPRIIGVWLFELIIGNALVI 352
QY 353 LTVFAGGAPLPPVKGAVGAIAGANTLTGISCGLASVDALTFQGFSEYGARWETGLCCR 412
DB 577 FVIIVSHKMDV-REFLCNLAFADEFGLGYLGLAGVDSTLGVFRFGKQWLSAGCR 635
QY 413 ATGFLAVLGSASVLLTLAAVCCSVSCVRAVGSPLGSVAGVGLA--LAGLAA 470
DB 636 LAGFLAVFSSEFSYITLSTVITLERFAIKHAKHLEKMKL--PHATVMEGWFISVTAA 693
QY 471 ALPLASVGEASPLCLPYAPPEGAPALGFTVALVMNSFCFLVAGAVIKYICDIPRG 530
DB 694 VLPLVNSHHRVPLCPF--DVDITVAKYVGSIIILNLITLAVIIMACASITLAI-QG 750
QY 531 DFEKVMDC--AMVRHVAWLIFADGLYCPVAFSPASMLGEPVPEAVKSVLLVLP 587
DB 751 SH--ANNCNDSRVARRMSLIVFTDFACMAPIAFFSLTAAGLKLISLDKAKVLLTVP 808
QY 588 PACLNLLVLLFNPFRDRLRL 610
DB 809 NSCANFELYTLIKOKKDKCKTI 831

RESULT 8
09DCG6
AC 09DCG6 PRELIMINARY; PRT; 701 AA.
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
GN Gonadotropin receptor 1.
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphia; Acanthopterygii; Perciformes; Labroidae;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=6128;
RN [1]
RP SEQUENCE FROM N.A.
RA Oba Y., Hirai T., Yoshiura Y., Yao Z., Nagahama Y.:
RT "Tilapia gonadotropin receptor 1."
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB041762; BAB16106.1;
DR HSSP: P23945; 1XUN.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR001611; LRR.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00560; LRR; 4.
DR PROSITE: PS00237; G_PROTEIN_RECPT_F1_1; UNKNOWN_1.
DR PROSITE: PS00262; G_PROTEIN_RECPT_F1_2; 1.
KW Receptor.
SQ SEQUENCE 701 AA; 78453 MW; B3D78465CA56410A CRC64;

```

```

Query Match 14.2% Score 545.5; DB 13; Length 701;
Best Local Similarity 26.8%; Pred. No. 4.7e-34;
Matches 175; Conservative 111; Mismatches 306; Indels 61; Gaps 20;

QY 2 LHNLETLIDNINKIOEFPV-AIRTLGRLOEGFNNN-IKAIPEKAPGNLLOTHTPYD 59
DB 59 ISMACCEVKQOIREIQOGLTSLLOHMLTISENDSIGAFASGPHLTKILISK 118
QY 60 N-PIQFVRSAPFOYLPKHTLSINGANDIQEF-PD-LKGTTSLEILLTRAGIRLPLSGM 116
DB 119 NAALRNIGAFEFNSULPELSEITITKSKHSLFHPDAFRNNAARLFLISMTGRIIPDES 178
QY 117 COQPLRVLVLELSN-OIELPS---LHRC-OKLEITGLHNRIRWEIGATFSQSLQA 171
DB 179 KIHSTACFLDLQDNSHIKRPANAFRGCTQTFPAERLRLRNQIKVASDAFNG-TKMR 237
QY 172 LDISSN-AIRSHPEAFSTLASHYKLDITDQTLPLAGLGLMHLTKGNLALSOAFS 230
DB 238 LFIQGNQQLHISPNAFVGSSELVLDVSETALTSLPSDITDLKRLIESAFNKLKELPP 297
QY 231 KDSFPRKILEVPAVOCCPYGMCAFFKASQMEADLHLDDESSKRLGLLAROEN 290
DB 298 IQPLTKLHQAKITYPHSCAF--LMHNRNRSRWS---LQDNPAKNNLHFFREYCSN 350
QY 291 HYDQDLDELQLEMDSKPPSVOCSPTPGPKPCEYLFESMGIRLAVMAIVLLSVLCNGL 350
DB 351 -----STNITCSPADPDPNCPDIMSAPPRLIITISVLLALGNAY 392
QY 351 VLTVPAGGAPLPPVKGAVGAIAGANTLTGISCGLASVDALTFQGFSEYGARWETGLG 410
DB 393 VLL-VLLSRKRLTVPRFLMCHLAFADLCMGILYIVATVMDLGRGRYNAIDMOMGLG 451
QY 411 CRATGFLAVLGSASVLLTLAAVCCSVSCVRAVGSPLGSVAGVGLA----- 465
DB 452 CNAAGFFTFVFASELSEVFTLTAIVERMHTTHAIRLDRKRLRH-----ACILMTIGWI 505
QY 466 -AGLAAALPLASVGEASPLCLPYAPPEGAPALGFTVALVMNSFCFLVAGAVIKY 524
DB 506 FSLAALLPTVGISSYKVSICLPM--DVESIVSQFVYVCLLNLITLAFVCGCYLSTY 563
QY 525 CDLPGRDFEAW-DCAVRRHVAWLIFADGLYCPVAFSPASMLGEPVPEAVKSVLLV 583
DB 564 LTRKPSAAAHADTRVAGRMALVIFTDFICAPISFFAISALKLDPLTVSDSKLLVL 623
QY 584 VLPLPACLNLLVLLFNPFRDRLRLPRPGDSGLAYAAAGEL---EKSSC 633
DB 624 FYIINGCSNFFLYAFTFRFRDFFLLAARFG-----LFKTRAGIYRTSSSC 671

RESULT 9
064183
AC 064183 PRELIMINARY; PRT; 688 AA.
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
OS Follicle-stimulating hormone receptor.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE FROM N.A.
RA Heckert L.L., Daley I.J., Griswold M.D.:
RT "Structural organization of the follicle-stimulating hormone receptor
gene."
RL Mol. Endocrinol. 6:70-80(1992).
DR EMBL: S81198; AAB21415.2; JOINED.
DR EMBL: S81117; AAB21415.2; JOINED.
DR EMBL: S81119; AAB21415.2; JOINED.
DR EMBL: S81121; AAB21415.2; JOINED.

```


DR EMBL: S81171; AAB21415.2; JOINED.
DR EMBL: S81174; AAB21415.2; JOINED.
DR EMBL: S81178; AAB21415.2; JOINED.
DR EMBL: S81183; AAB21415.2; JOINED.
DR EMBL: S81185; AAB21415.2; JOINED.
DR EMBL: S81194; AAB21415.2; JOINED.
DR HSSP: P23945; 1XUN.
DR InterPro: IPR000276; GPCR_Rhodopsin.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000372; LRR_Nterm.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PF00560; LRR; 3.
DR Pfam: PF01462; LRRNT; 1.
DR SMART: SM00013; LRRNT; 1.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR Receptor.
SQ SEQUENCE 688 AA; 77341 MW; 441F0D9E7D01DF18 CRC64;

Query Match 13.1%; Score 504.5; DB 11; Length 688;
Best Local Similarity 25.7%; Pred. No. 7,4e-31;
Matches 186; Conservative 106; Mismatches 283; Indels 149; Gaps 25;

QY 30 ELGFHNNNIKAIPKAFMGNPLLOTIHFYDNPQFVGRSAFOYLPKRLHTLSNGAMDIO 88
DB 50 ELRFVLTKLRLVIRPKGSFAGGDELEKIEISONDYLEVLEADVFSNLPKHLRIEKKANL- 108
QY 89 EPPDLKGTTSLEILLTRAGIRLLPSGMCQOLPRLRVLELSHNOIEELPSLHRCQKLEI 148
DB 109 -----LYINPEAF-QNLPRLKRLLSNTGIRKHLPAVHKIQSLQKV 147
QY 149 --GLOHN-RIMEIGADTFSSLSQALDLSMNAIRSIHPEAF-----STLHS 192
DB 148 LLDIDQININIHVARSEFMSFEW--LSKNGIEEIHNCANFGTODLDELNSDNNNEE 204
QY 193 L-----VKLDLTQDLTTLPLAGLGMLHLKLGMLALSOAFSKDSPKRLILE 241
DB 205 LPRDVFQASGPVILDISRKVSHSLPHNGLENLKKLRARSTYRKAKLPNDKFTLMAS 264
QY 242 VRYAYOCPCPYGMCASFEGKASQWEAEDLH-----LDEES 276
DB 265 LTVPSHCCAF-----ANLKRQISELHPICNKSILRQIDDMTOIGDQVSLIDDEPS 316
QY 277 SKRPLGLARQAEHNDQDLDELQLEMEDSKPHPSVOCSPTPGPKCEYLFESNGIRLA 336
DB 317 ---YKGSDDMYNEPDYDCN---EYVD-----VTCSPKPDAPNCEDEIMGYNIIRVL 363
QY 337 VMAIVLTVLCNGLVLTAVFAGPAPLPYKFPVYGATAGANTLGLISGGLASVDALTFEG 396
DB 364 IMFISLILATIGNTTIVL-VLTTSQYKLTVPREFIMCNLAFAADLCIGIYLLIASVDIHTKS 422
QY 397 QPSEYGARWETGLGCRATGFLAVLGEASVLLTLAAVO-----CSYVSQVAYGKSP 450
DB 423 QYHNVAIDMOTGAGCDAAGFTVAFASLSTYTLAIIERHNTTHAMQLEC-----KVQ 477
QY 451 SLGSVAGVLCALAGLAALPLASVGEYASPLCPYAPREGQAPALGFTVALVMNS 510
DB 478 LRRHAAVMYLG-WTFAFAALFPFIIGISSYMKVSICLPM--DIDPSQLQTLVMMLLVIN 534
QY 511 PCELVAVAGVYIKYCDLPRDF-EAVWDCAMVRHVAWLIFADGLLYCPVAFLSASMGIL 569
DB 535 LAFVYICGCTTHIYLVNRNPTIYSSSDTIKAKMATLIFDFLCMAPISEFFAISASLKV 594
QY 570 PFTVPAVKSLLVLPPLPACLPPLLYLTFNPHFRDRLRLRPRAAGSGPLAAYAAGLE 629
DB 595 PLTVSKAKILLVLPFINSKANPLIYFTKNFRDFIL-----LS 637
QY 630 KSSCDSTQALVAFSDVDLLIEASAGRPCELETYGFPSVLLISQGPAGAPLESCHVEP 689
DB 638 KFGCYEQMOAIYRT-----ETSSA-----THNF-HARKSHCS--SAPRTVNSYLVLP 681
QY 690 EGNH 693
II

DB 682 L-NH 684
RESULT 10
Q8R428 PRELIMINARY; PRT; 695 AA.
AC Q8R428;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Follicle stimulating hormone receptor.
GN FSHR.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.
OX NCBI_Taxid=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HARTLEY; TISSUE=TESTIS;
RA Suzuki O.;
RT "Guinea pig follicle stimulating hormone receptor."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY082514; AAL92577.1; -
KW Receptor.
SQ SEQUENCE 695 AA; 77838 MW; 9A3ECF419C45227B CRC64;

Query Match 12.7%; Score 490.5; DB 11; Length 695;
Best Local Similarity 25.3%; Pred. No. 9,3e-30;
Matches 166; Conservative 111; Mismatches 286; Indels 93; Gaps 19;

QY 30 ELGFHNNNIKAIPKAFMGNPLLOTIHFYDNPQFVGRSAFOYLPKRLHTLSNGAMDIO 88
DB 50 ELRFVLTKLRLVIRPKGSFAGGDELEKIEISONDYLEVLEADVFSNLPKHLRIEKKANL- 108
QY 89 EPPDLKGTTSLEILLTRAGIRLLPSGMCQOLPRLRVLELSHNOIEELPSLHRCQKLEI 148
DB 109 -----LYINPEAF-QNLPRLKRLLSNTGIRKHLPAVHKIQSLQKV 147
QY 149 --GLOHN-RIMEIGADTFSSLSQALDLSMNAIRSIHPEAF-----TLH 191
DB 148 LLDIDQININIHVERNSFGLSESVLLRLKNGIQEIQCAFNFGTODLDELNSDNNLE 207
QY 193 L-----VKLDLTQDLTTLPLAGLGMLHLKLGMLALSOAFSKDSPKRLIL 240
DB 205 LPRDVFQASGPVILDISRKVSHSLPHNGLENLKKLRARSTYRKAKLPNDKFTLMAS 264
QY 242 VRYAYOCPCPYGMCASFEGKASQWEAEDLH-----LDEESKRPGLT---LARQAE 289
DB 265 LTVPSHCCAF-----ANMRQISELHPICNKSILRQEVNDITQAGKORVSLADE 319
QY 277 SKRPLGLARQAEHNDQDLDELQLEMEDSKPHPSVOCSPTPGPKCEYLFESNGIRLA 336
DB 317 ---YKGSDDMYNEPDYDCN---EYVD-----VTCSPKPDAPNCEDEIMGYNIIRVL 363
QY 337 VMAIVLTVLCNGLVLTAVFAGPAPLPYKFPVYGATAGANTLGLISGGLASVDALTFEG 396
DB 364 IMFISLILATIGNTTIVL-VLTTSQYKLTVPREFIMCNLAFAADLCIGIYLLIASVDIHTKS 422
QY 397 QPSEYGARWETGLGCRATGFLAVLGEASVLLTLAAVO-----CSYVSQVAYGKSP 450
DB 423 QYHNVAIDMOTGAGCDAAGFTVAFASLSTYTLAIIERHNTTHAMQLEC-----KVQ 477
QY 451 SLGSVAGVLCALAGLAALPLASVGEYASPLCPYAPREGQAPALGFTVALVMNS 510
DB 478 LRRHAAVMYLG-WTFAFAALFPFIIGISSYMKVSICLPM--DIDPSQLQTLVMMLLVIN 534
QY 511 PCELVAVAGVYIKYCDLPRDF-EAVWDCAMVRHVAWLIFADGLLYCPVAFLSASMGIL 569
DB 535 LAFVYICGCTTHIYLVNRNPTIYSSSDTIKAKMATLIFDFLCMAPISEFFAISASLKV 594
QY 570 PFTVPAVKSLLVLPPLPACLPPLLYLTFNPHFRDRLRLRPRAAGSGPLAAYAAGLE 629
DB 595 PLTVSKAKILLVLPFINSKANPLIYFTKNFRDFIL-----LS 637
QY 630 KSSCDSTQALVAFSDVDLLIEASAGRPCELETYGFPSVLLISQGPAGAPLESCHVEP 689
DB 638 KFGCYEQMOAIYRT-----ETSSA-----THNF-HARKSHCS--SAPRTVNSYLVLP 681
QY 690 EGNH 693
II

RESULT 11

```

O9PVP0 PRELIMINARY; PRT; 724 AA.
AC O9PVP0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Gonadotropin receptor II.
GN SGT-II.
OS Oncorhynchus rhodurus (amago).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=41164;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99423507; Pubmed=10491336;
RA Oba Y., Hirai T., Yoshikuni M., Kawauchi H., Nagahama Y.;
RT "Cloning, functional characterization, and expression of a
gonadotropin receptor cDNA in the ovary and testis of amago salmon
(Oncorhynchus rhodurus).";
RL Biochem. Biophys. Res. Commun. 263:584-590(1999).
DR EMBL; AB030005; BAA84638.1; -.
DR HSSP; P22888; 1LUT.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSN.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECPEP_FL_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECPEP_FL_2; 1.
KW Receptor.
SQ SEQUENCE 724 AA; 80494 MW; E999652740332B2E CRC64;

Query Match 12.6%; Score 487; DB 13; Length 724;
Best Local Similarity 24.7%; Pred. No. 1.8e-29;
Matches 165; Conservative 113; Mismatches 262; Indels 128; Gaps 21;

OY 1 GLNHLTLDLNYKNLQEPVPAIRTLGRLOELGFHNNNIKAIEKAFMGNPDLQTHIFYDN 60
DB 76 GLRRVQHEIGQS-----VALET--LETIAF--NNLUDLNE-IFIKN-IRSLVH--- 118
OY 61 PLOFVGRSAFOYLRLKLTLSINGAMIDIOEPDLKGTSTSE-----ILVTRAGIRLLPS 114
DB 119 -----IARTFNMLPKRLRLTSISNT-GITVFPRMSTSIHSLPPNNQNVLDI----- 163
OY 115 GMCOQLRPLRV-----LELSHNOIEELPS-LHRCQKLEETIGLOHNRIMEIG 159
DB 164 --CDNLVLTLSIPVNAFVGMTTEYTAAMNFNGNIREIDYAFNGTKINKLVKNNR----- 216
OY 160 ADFEQLSSIQALDLSNMATRSIHPEAFSTLSLYKLDLTDLQTLTLPRLAGLGGLHMLK 219
DB 217 -----NLRVIHREAFKGAAGPRILDDVSTAIEFLTPSHGLNSVVELVA 258
OY 220 KGNLALSOAFSKDSFPKRLRIEVPYAYQCCPYGMCASFPAKASGOME----- 265
DB 259 RTAYGLKRLPFRRLGNLQKRLHLYNSHC-----ALLTWDTHADSPINEAQHNG 308
OY 266 AEDLHLDEESSKRPLGL-----ARQAEHHYDODLDELQLEMEDSKPHPSV 312
DB 309 SRPLFYCDSPSDKFRPAGVDSSTSLVEIHGTNEDESVGGVDFQPELML-NCQTRPPL 367
OY 313 QCSPTPEPFPCETLFSWGIRLAWAIVLLSVLCNGVLLTVFVAGGAPRLDPVKVGVCA 372
DB 368 QCPEADAFNPCEIDIAFSSFLRAIWFNIIILATION-LTVLLIFTSRCKLVAPREFLMCH 426
OY 373 IAGANTLTGISCGLASVDALTFGQFSEYGAARMWENGLGRATGGLAVLISEASVLLTLTA 432
DB 427 LARADFCIGYLLMTAIVDLHTTRGHYSHAIDMCRGAGSAGAFUSVFGGELSVYTLSTI 486
OY 433 AVQ-----CSVSVSCVRAVYGKSPSLGVSAGVLCIALAGLAALPLASVGEYGASPLCLP 488

```

RESULT 12

```

O9PVP0 PRELIMINARY; PRT; 658 AA.
AC O9PVP0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Gonadotropin receptor I.
GN SGT-I.
OS Oncorhynchus rhodurus (amago).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=41164;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20025349; Pubmed=10558873;
RA Oba Y., Hirai T., Yoshikuni M., Kawauchi H., Nagahama Y.;
RT "The quality of fish gonadotropin receptors: cloning and functional
characterization of a second gonadotropin receptor cDNA expressed in
the ovary and testis of amago salmon (Oncorhynchus rhodurus).";
RL Biochem. Biophys. Res. Commun. 265:366-371(1999).
DR EMBL; AB030012; BAA86898.1; -.
DR HSSP; P22888; 1LUT.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR001611; LRR.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 2.
DR PRINTS; PR00237; GPCR_RHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECPEP_FL_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECPEP_FL_2; 1.
KW Receptor.
SQ SEQUENCE 658 AA; 72991 MW; 853A38BEA845480 CRC64;

Query Match 12.5%; Score 482.5; DB 13; Length 658;
Best Local Similarity 26.4%; Pred. No. 3.6e-29;
Matches 161; Conservative 100; Mismatches 285; Indels 63; Gaps 14;

OY 30 ELGFHNNNIKAIEKAFMGNPDLQTHIFYDN-PIQFVGRSAFOYLRLKLTLSINGAMDI- 87
DB 56 DLEFGQTHIRVPRQEFNLTQDLTAVLVEHNGMLESIGAFAPANLPLRTEITTKSHLV 115
OY 88 -----QEPDLKGTSTSEITLTRAGIRLLPS-GMCOQLRPLRVLELSHN-QIEELPSIH- 140
DB 116 IHHQAF--IDLPRKSHLTICNTGLRVLPNFSRISHAAMTFLLDODNVIIIVIPSNVAF 172
OY 141 ---RQKLEETIGLONHRIWEIGADPFSQSLQALDSMN-AIRSIHPEAFSTLSLYKL 196
DB 173 LGLTNTTIDELRLTKNGSIEVESHAENG-TKIKHLYMGLOLSHMINNSKGAEGEGL 231
OY 197 DLTDLQTLTLPRLAGLGMLKLMKGNLALSOAFSKDSFPKRLRIEVPYAYQCC----- 249
DB 232 DISKRLALSLPSVGLVEYHLSAVSFSLRTEPLPLFTLRLQGANLTPPSGCCAFHHKQR 291
OY 250 --PYGMCASFPAKASGOMEADLHLDEESSKRPLGLIARQAEHHYDODLDELQLEMEDSK 307
DB 292 NRTFRMTSACFKPQAO-----NNLHFFMDFCL 318

```

QY 308 PHSVQSGPTGPFKRCCEYLFEFSGIRLAVMAIVLISVLCNGILVLTTRAGCPAPLPVK 367
 319 NMTSVACSPAPDANPCEDIMGSAFLRLIWIISVLLALGNITVL-VLLSGRAKMTVR 377
 QY 368 FVVGAIAGANTLGLISGGLASVDALTPOGFSYXGAWMETGLGCRATGFLAVGSAVYL 427
 378 FLMCHLSFADLCMGILVAVIATVDTRGLYINHAISWOTGACCDIAGFTTVASELSMF 437
 QY 428 LTLIAVQGSVSVCAVYKSPSGSVRAGVGLCLAGLAAALPLASVGEYASPLCL 487
 438 TLTATLTERKHTHTHRLDRKRLRHACAVMATGMAFSCIALALPTVGVSSYSKYSICL 497
 QY 488 PVAPPEGAPALGETVALVYMMNSFCLVAVAGATIKLYCDLPBGDF-EAVYDCAMVRIVAN 546
 498 P-MDVESTPSQV-FVMFLLLINVAFLCYVCYLSYLSVSNSSPPASAEQMAQMAI 555
 QY 547 LTFADGLICPVAFLEPSASMLGLEPVTPEAVKSVLLVPLPACLNPLTLLENPHRRD 606
 556 LITTDLCMAPISEFALSALKPLITVSDSKLLVLFYFINSCANPFLXGICTRFRD 615
 QY 607 LRLRPBAG 615
 DB 616 FFLAARYG 624

RESULT 13
 QY18N7 PRELIMINARY; PRT; 779 AA.
 AC 0918N7;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Thyrotropin receptor precursor.
 OS Morone saxatilis (Striped bass).
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
 OC Moronidae; Morone.
 NCBI_Taxid=34816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=OVARY;
 RX MEDLINE=20457109; PubMed=11000515;
 RA Kumar R.S., Jirli S., Kight K., Swanson P., Diltman A., Alok D.,
 RA Zohar Y., Trent J.M.;
 RT "Cloning and functional expression of a thyrotropin receptor from the
 RT gonads of a vertebrate (bony fish): potential thyroid-independent role
 RT for thyrotropin in reproduction.";
 RL Mol. Cell. Endocrinol. 167:1-9(2000).
 DR EMBL: AF239761; AAF80596.1; -.
 DR HSSP: P16473; 1XUM.
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR InterPro: IPR001611; LRR.
 DR Pfam: PF00560; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 2.
 DR PRINTS: PR00237; GPCR_RHODOPSN.
 DR PROSITE: PS00237; G_PROTEIN_RECPT_F1_1; UNKNOWN_1.
 DR PROSITE: PS0262; G_PROTEIN_RECPT_F1_2; 1.
 KW Receptor; Signal.
 FT SIGNAL 1
 FT SEQUENCE 779 AA; 87400 MW; 8C86D121B676A9EF CRC64;

Query Match 12.4%; Score 478.5; DB 13; Length 779;
 Best Local Similarity 23.5%; Pred. No. 9,4e-29;
 Matches 177; Conservative 118; Mismatches 315; Indels 143; Gaps 21;
 QY 3 HNLFTLDLNYKQEPVAIRLGRLOELGFHNNNKAIPKAFKGNPLQITLHF-YDNP 61
 DB 38 HTISCDDIDI-LPRPASTETLMLFE-----TSLSSVPADAFSSWMNSIRIYISVDMT 89
 QY 62 IQFVGSARQYLPKLTLSLNGAMDIOEPFDLKGTSLELTTLTRAGIRLLPSCMGCOLP 121

DB 90 LQRLERHSFYSLEKRTIHTIIRNA-----KSLTYID-----PEAF-KNLP 127
 QY 122 RLRYLELSHQIETELPSLHRCCK-----LEEIGLOHNRIMELGADTFEOLS-LOALDLS 175
 DB 128 NLKYLIGFNGIIFFPDLNITSNDNFLEI-VDHYTEIETPSRIGTSYLVFLVLM 186
 QY 176 WNAIRSIHPEAFS-----TLHSYKLDLDTNQLTTLPLAG 210
 DB 187 GNGFRELQHNHAFNGTKLDQVDLHRNKYLTMRDERAFAGTISGMLDLVSLTGITLPTTG 246
 QY 211 LGLIMHLKLGKLNALSAFESKDSFPRKRIIEVPAVQCC-----PYGMC--AS 256
 DB 247 MDSIRELKARNAMALKPLPKTFKHLTIANLTYPRHCCGFKMLKKRGLEYITICMLTA 306
 QY 257 FF-----KASGQEAEDLH----- 270
 DB 307 FYDOHNRKSVGLRIPSLQGESVETIPDQEPNDGHRSGQDRDGFHSLHYHAYFG 366
 QY 271 -LDDE-----ESSKRPVGLLARKAENHYDQDLDELQLEMEDSKRPHSVQSGPTGPKPC 324
 DB 367 GQPEDVGFGEFTLKNPQEDTSQDPDSRDYVCE--EGEE-----VSCAPVEDERNPC 417
 QY 325 EYLFESNGIRLAVMAIVLISVLCNGILVLTTRAGCPAPLPVKFVVGAIAGANTLGLISC 384
 DB 418 EDINGFGFLRVSVFVSLAVLGNVYVL-VLISHYKLSYRFLMCHLAFADLCMGITYL 476
 QY 385 GLIASVDALTFPGFSEYGARWETGLGCRATGFLAVLGSASVLLITLTAVQGSVSVQVR 444
 DB 477 LLIASVDLHTAEYFNHAIIDQGTGGCLAGFTVYFASLSVYTLVTLTTRWYATFTAM 536
 QY 445 AYKSPSISGSRVAGVGLCLAGLAAALPLASVGEYASPLCLPAPPEGAPALGETVA 504
 DB 537 RLDRKLHHAAYVLLGWITFCLLALPLVGVSSYQKVSICLPR-DTQSVNAVYILS 594
 QY 505 LYMMNSFCLVAVAGATIKLYCDLPBGDEA-WMDCAMVRIVANLTFADGLICPVAFISF 563
 DB 555 VLVNIIILAFVVICAFYFYCAVHNPHYRSGSKDQINIAKRMALVLTFTFLCAPISFYAM 654
 QY 564 ASMTGLFPVTPKAVSVLLVPLPACLNPLTLLENPHRRDLRLRPBAGDSPLVYA 623
 DB 655 SAVIDRPLITVSNKLTLLVLFPLNSCANPELYLFTAFRGDVFILLSKVLCOQRAL 714
 QY 624 AAGELEKSSCDSTOALVAFSDVDLILEASEAGR 656
 DB 715 FRGQIVSSKSGSSGTQVR-RQDKVRKSGSGQ 746

RESULT 14
 QY18N7 PRELIMINARY; PRT; 693 AA.
 AC 09DGC5;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Gonadotropin receptor II.
 OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostei;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
 OC Cichlidae; Oreochromis.
 NCBI_Taxid=8128;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Oda Y., Hirai T., Yoshitani Y., Yao Z., Nagahama Y.;
 RT "Tilapia gonadotropin receptor II.";
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB041763; BAB16107.1; -.
 DR HSSP: P22888; 1LUT.
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR InterPro: IPR001611; LRR.
 DR Pfam: PF00560; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 2.

```

DR PRINTS; PRO0237; GPCRHDOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECPT_FL_1; UNKNOWN_1.
DR PROSITE; PS00262; G_PROTEIN_RECPT_FL_2; 1.
KW Receptor.
SQ SEQUENCE 693 AA; 77856 MW; C3BBEFDCEFC8988 CMC64;

Query Match
Best Local Similarity 27.7%; Score 475.5; DB 13; Length 693;
Matches 163; Conservative 98; Mismatches 237; Indels 91; Gaps 21;

OY 60 NPQVGRSAFOYLPRKHTLSLNCAMDIOEPDRLKGTSLSLITLITRAGIRLSPGCMQ 119
DB 112 NSLRFLFKGAFADLPKREYLSISWT-GIAHPDPTTSSLS----- 151
OY 120 LPLRLVLESHN-OIEELPSLHRCQKLEE---IGLOHNRIMEIGADTF--SOLSLQAL 172
DB 152 -PNI-ILEMADNEIDPIA-NSFOGITEEVDMNLVRNGEKSHAFNCTKNTLVL 208
OY 173 DLSMNAIRSIHPAFSTLSLVLKIDLDNOLITPLAGLGMLKLGKLNALSOAFSKD 232
DB 209 D-NW-YLRNIQEDAFEGAGPTLIDVSTALRSLP---PNGLRHVKE---LKASHAYALK 260
OY 233 SEPKIRIL-----EYVAYQCCPYGMCASFYKASQGMEDLHLDDESSKRLGLLAR 286
DB 261 SLPLESLAELEAEELTPSHCC-----AFHWRRK-----QRESALKLITFC 305
OY 287 QAEHNHDDLDLQI-----EME-DSKPHPSVQCCPTPGFKPCFVLFESWGIRLA 336
DB 306 LMNTEIDTPADTSLINDINFOYRPLEDFDSCNPFVYCSKPKPDAFNCEIDLGRSFLRCL 365
OY - 337 WVAIVLVSICNLVILTFVAGAPAPLPKPVVGAAGANTLTGICGGLASVADALTFG 396
DB 366 TWIMVFAVAGN-LAVLVLLIGHNKLTVSRFLMCNLAFLDLCGLYLILIAFDYSHH 424
OY 397 OFSEGARRETLGCRGATFLAVLGSEASVLLTLTAAQCVSVCRAVGSFSLSVR 456
DB 425 EYNNATDMQTPGCGIAGFLTFVSELSVYTLFTVLSERWHTITNAMHNKRLMHHTV 484
OY 457 AGVLGCLLAGLAALPLASVGEYGSPLCLPAPRPGOPALG---FTYALVMNSFCF 513
DB 485 AMNVGMAFSLVALLPLGVSSYSKVSICLPV---DIDTGAOVYVAVLILNVAF 539
OY 514 LVVAGAYIKLYCDL-----PRGDFEAVWDCAMRVHAMLFPADGLICPVAFPSASM 566
DB 540 LVVCCYICILYLVNHPHSTRGDK-----IAKMAVLITFDLCAPIFFAISAA 593
OY 567 LGLEPVPEAVSVLVVLPACINPLLYLFPNPHFRODLRLRPRAG 615
DB 594 LRMPLITVSHSKILILITPLINSICNPFLYITFTFRAFRKDVCLLSRCG 642

RESULT 15
OYVEG4 PRELIMINARY; PRT: 829 AA.
AC OYVEG4;
DT 01-MAY-2000 (TREMUREL.13, Created)
DT 01-MAY-2000 (TREMUREL.13, Last sequence update)
DT 01-MAR-2002 (TREMUREL.20, Last annotation update)
DE FSH protein.
DE FSH OR CG7665.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Cealiker S.E., Holt R.A., Evans C.A., Coeayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

```

```

RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abill J.F., Agapayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Glodex A., Gong F., Gorrell J.H., Garg N.S., Gelbart W.M., Glasser K.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Mostrel A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclebo J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zhang X.H., Zhong F.W., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AE003719; AAF5460.1; -.
DR FlyBase; FBgn0016650; Fsh.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR001611; LRR.
DR Pfam; PR00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 2.
DR PRINTS; PRO0237; GPCRHDOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECPT_FL_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECPT_FL_2; 1.
KW G-protein coupled receptor; Glycoprotein; Transmembrane.
SQ SEQUENCE 829 AA; 92583 MW; 2224C38910D77856 CMC64;

Query Match
Best Local Similarity 22.3%; Score 475.5; DB 5; Length 829;
Matches 158; Conservative 129; Mismatches 275; Indels 145; Gaps 19;

OY 1 GLHNL-ETLDLNVNKLDEFPVARTLGLQELGHNH-----NIKARPEAF 46
DB 120 GUNRVQTLT---PIORLITASGLRLRTGTLKAVGSLTLDVAFTDCLQLELDGDAF 176
OY 47 MGNPLDITHFYDNP-IQVGRSAFOYLPRKHTLSLNGAMDIOEPDLKGTSLSLITLT 105
DB 177 ANTLTLRTIYTNAKPLTFLSKDVF-----LGISDTVDI----- 210
OY 106 RAGIRLPSGMCQ-----QLPR---LVLELSHQIIEELPSLHRCQKLEEGLOHNRIME 157
DB 211 ---IRIINSLTVPDGLPHNIILOMIDLDNNOIRIRDSIKVATQILITNNEISY 267
OY 158 IGADTF--SOLSSLOALDLSMNAIRSIHPAFSTLSLVLKIDLDNOLITPLAGLGIM 215
DB 268 VDSAFPGSKIAKQ-----ENKRLQMMHNPALFGIIDIITELDLSSTLVGPSAGIONIE 323
OY 216 HLKLGKGNLALSOAFSKDSEPKIRILEVPVAYQCC----- 249
DB 324 ALYIONTHLTKTPTSIYNFRNIIQRAVLTHSFHCACAFQPSRHDPORNAQMLETEKRRKQ 383
OY 250 -----PYMCASFYKASQGMEDLHLDDESSKRLGLLARQAEHNHYD 293

```

```
Db 384 CKSDSGTRENSTLNPNMPEDEFGSG-----TDDSATDITPIFFAS-----FD 429
QY 294 QDLDELQLE-----MEDSKPH-----PSVQCSPTGEPKPCCEYLEESWGI 333
Db 430 YMADDTMNKGTFHEKILNPGDSSAELCGNFTFRKPNIECYMPNDLNPCEDVMGYQWL 489
QY 334 RLAVNAIYLLSYLNGVLVLYFAGGPAFLPVKFFVGAAGANTLTGISCGLLASVDL 393
Db 490 RISWIVVALAVGNAVAVLTVLISIRPESTVPVPRFLMCHLAFADLCLGLYLLLVACIDAH 549
QY 394 TFGOFSEYGARMTGIGCRATGFLAVLGSEASYLLTLTAVOCSVSCVRAYGKSPSLG 453
Db 550 SMGEYFNFAIDMVGIGCKVAGFLTYFASHLSVFTLTITIERMLAITQAMYLNHRILR 609
QY 454 SVRAGVLGLALAGLAALPLASVGEYGASPLCLPYAPPEGOPALGFTVALYMMNSFCF 513
Db 610 PALIMLGMIYSMLMSSLPFGISNYSSTISICLPENRDVYDTI--YLAILGNGVAF 667
QY 514 LVVAGAYIKLYCDLPRGDPEA---VWDCAMVRHVAMLIFFADGLLYCPVAFLSFASMLGL 569
Db 668 SIIAVCYAQIYLSLGRHETQAHONSPGELSVAKKMLLVFTNFACWSPIAFGLTALAG- 726
QY 570 FPVTPKAVKSVLLVVL-PLPACINPLLYLLFNPHRDDLRRLRPAG 615
Db 727 YPLINVTKSKILLVFYPLNSCADPYLYAILTSQYRODLFTLLSKLG 773
```

Search completed: November 8, 2002, 19:36:17
Job time : 43.5581 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2002, 19:31:55 ; Search time 10.7524 Seconds
(without alignments)
2839.055 Million cell updates/sec

Title: US-09-851-595-8

Perfect score: 3850
Sequence: 1 GLHNLERLDLNYNKLQEPV.....GGLSGGGGFGPFGIAFASHV 736

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3818	99.2	828	1	LGR6_HUMAN
2	1844	47.9	907	1	LGR5_HUMAN
3	1780	46.2	907	1	LGR3_MOUSE
4	1410	36.6	951	1	LGR4_RAT
5	1409	36.6	951	1	LGR4_HUMAN
6	526.5	13.7	695	1	FSHR_BOVIN
7	524.5	13.6	695	1	FSHR_SHEEP
8	523	13.6	695	1	FSHR_PIG
9	522	13.6	694	1	FSHR_HORSE
10	521.5	13.5	692	1	FSHR_MOUSE
11	513.5	13.3	692	1	FSHR_RAT
12	502.5	13.1	695	1	FSHR_MACFA
13	498.5	12.9	695	1	FSHR_HUMAN
14	492.5	12.8	687	1	FSHR_EOVAR
15	463	12.0	763	1	FSHR_BOVIN
16	460	11.9	764	1	FSHR_CANFA
17	457.5	11.9	693	1	FSHR_CHICK
18	456.5	11.9	696	1	FSHR_PIG
19	456.5	11.9	764	1	FSHR_SHEEP
20	454.5	11.8	764	1	FSHR_MOUSE
21	449.5	11.7	700	1	FSHR_MOUSE
22	446	11.7	701	1	FSHR_BOVIN
23	443.5	11.6	700	1	FSHR_RAT
24	436	11.3	764	1	FSHR_HUMAN
25	425.5	11.1	764	1	FSHR_HUMAN
26	424.5	11.0	764	1	FSHR_RAT
27	421.5	10.9	737	1	LGR8_MOUSE
28	413	10.7	699	1	FSHR_HUMAN
29	411.5	10.7	676	1	FSHR_CALJA
30	394	10.2	538	1	FSHR_SHEEP
31	389	10.1	757	1	LGR7_HUMAN
32	384	10.0	925	1	GLHR_ANCEL
33	357	9.3	366	1	LSHR_CHICK

34	338	8.8	1115	1	GPCR_LYMST
35	233.5	6.1	536	1	GPB8_HUMAN
36	233	6.1	331	1	PUB_AGRBL
37	228	5.9	582	1	SHO2_HUMAN
38	227.5	5.9	603	1	ALS_MOUSE
39	226	5.9	605	1	ALS_HUMAN
40	226	5.9	605	1	ALS_PAPHA
41	224	5.8	582	1	SHO2_MOUSE
42	222.5	5.8	603	1	ALS_RAT
43	220	5.7	1039	1	YR71_CAEL
44	219.5	5.7	782	1	CHAO_TRICA
45	218	5.7	567	1	GPV_RAT

ALIGNMENTS

RESULT 1
LGR6_HUMAN
ID LGR6_HUMAN STANDARD: PRT: 828 AA.
AC Q9HXB8; Q96K69;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Leucine-rich repeat-containing G protein-coupled receptor 6.
GN LGR6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20388592; PubMed=10935549;
RA Hsu S.Y., Kudo M., Chen T., Nakabayashi K., Bhalla A.,
RA van der Spek P.J., van Duin M., Hsuen A.J.W.;
RT "The three subfamilies of leucine-rich repeat-containing G protein-
RT coupled receptors (LGR): identification of LGR6 and LGR7 and the
RT signaling mechanism for LGR7";
RT Mol. Endocrinol. 14:1257-1271(2000).
RN [2]
RP SEQUENCE OF 406-828 FROM N.A.
RX Takeda S., Kadowaki S., Haga T., Takaesu H., Mitaku S.;
RT "Identification of G protein-coupled receptor genes from the human
RT genome sequence.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 431-828 FROM N.A.
RX TISSUE-Mammary gland;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA Tanase T., Nomura Y., Togiya S., Komai F., Saito K., Yamamoto J.,
RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;
RT "NBD human cDNA sequencing project.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -i- FUNCTION: Orphan receptor.
CC -i- SUBCELLULAR LOCATION: Integral membrane protein.
CC -i- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -i- SIMILARITY: CONTAINS 11 LEUCINE-RICH REPEATS (LRR).
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF190501; AAC17168.1; -
DR EMBL: AB083616; BAB89329.1; -
DR EMBL: AK027377; BAB55071.1; ALT_INIT.
DR MIM: 606653; -
DR InterPro: IPR000276; GPCR_Rhodpsn.

```

DR InterPro; IPR001611; LRR.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00001; 7tm.1. 1.
DR Pfam; PF00560; LRR; 8.
DR PRINTS; PRO1561; EDG8RECEPTOR.
DR PRINTS; PRO0237; GPCRHHODOPSIN.
DR PRINTS; PRO0019; LEURICHRPT.
DR SMART; SM00370; LRR; 5.
DR SMART; SM00369; LRR_TYP; 10.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; FALSE_NEG.
DR G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat;
KW Leucine-rich repeat.
FT DOMAIN 1 428
FT TRANSMEM 429 449
FT DOMAIN 450 459
FT TRANSMEM 460 480
FT DOMAIN 481 505
FT TRANSMEM 506 526
FT DOMAIN 527 548
FT TRANSMEM 549 569
FT DOMAIN 570 588
FT TRANSMEM 589 609
FT DOMAIN 610 635
FT TRANSMEM 636 656
FT DOMAIN 657 670
FT TRANSMEM 671 691
FT DOMAIN 692 804
FT TRANSMEM 813 816
FT REPEAT 22 45
FT REPEAT 46 69
FT REPEAT 71 93
FT REPEAT 94 117
FT REPEAT 118 140
FT REPEAT 142 164
FT REPEAT 188 211
FT REPEAT 212 236
FT REPEAT 238 257
FT REPEAT 258 281
FT REPEAT 283 305
FT DISULFID 303 578
FT CARBOHYD 15 15
FT CARBOHYD 34 34
FT CONFLICT 406 410
FT CONFLICT 628 628
FT CONFLICT 824 824
SQ SEQUENCE 828 AA; 89301 MW; 1B5971445AA2D8B4 CRC64;

```

Query Match 99.2%; Score 3818; DB 1; Length 828;
 Best Local Similarity 99.3%; Pred. 3.2e-266;
 Matches 730; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

QY 2 LHNLETLDTLNKIQEPPVAIRTLGRLQELGFHNNNIKAIPKAFMGNDLQTIHFYDNP 61
DB 94 LPSLOSIDLNNKIQEPPVAIRTLGRLQELGFHNNNIKAIPKAFMGNDLQTIHFYDNP 153
QY 62 IOFGRSAFOYLPKLTHTLSNGAMDIQEPRLKGTSTLEILTRAGIRLPSGMCQQLP 121
DB 154 IOYVGRSAFOYLPKLTHTLSNGAMDIQEPRLKGTSTLEILTRAGIRLPSGMCQQLP 213
QY 122 RLNLVELSHNQIEELPSLINCQKLEETIGLQHNRIWEIGADTFSSQLSLOALDSNNATRS 181
DB 214 RLNLVELSHNQIEELPSLINCQKLEETIGLQHNRIWEIGADTFSSQLSLOALDSNNATRS 273
QY 182 IHEAPSTLSLYKLDLTNQTLTLPPLAGLGIMHLKLGMLALSOAFSKDSFKRLTLE 241
DB 274 IHEAPSTLSLYKLDLTNQTLTLPPLAGLGIMHLKLGMLALSOAFSKDSFKRLTLE 333
QY 242 VPAYAOCCPRGKMGASFPGKAGOWEADLHDDDESSKRPGLGLARQENHYDQDLDELQQL 301
DB 334 VPAYAOCCPRGKMGASFPGKAGOWEADLHDDDESSKRPGLGLARQENHYDQDLDELQQL 393

```

```

QY 302 EMEDSKPHPSVOCSPPTGPEPKPECEYLFESWGIRLAWAIVLLSYLNGVLTIVFAGGPA 361
DB 394 EMEDSKPHPSVOCSPPTGPEPKPECEYLFESWGIRLAWAIVLLSYLNGVLTIVFAGGPA 453
QY 362 PLPYKFEVVGATAGANTLIGISGGLASVDALTFEGQSEYGARWETGLGCRATGFLAVLG 421
DB 454 PLPYKFEVVGATAGANTLIGISGGLASVDALTFEGQSEYGARWETGLGCRATGFLAVLG 513
QY 422 SEASVLLTLTAAVOCSSVSVCAVAYGKSPSLGSRVAGVGLCLAGLAAALPLASVGEYG 481
DB 514 SEASVLLTLTAAVOCSSVSVCAVAYGKSPSLGSRVAGVGLCLAGLAAALPLASVGEYG 573
QY 482 ASPLCLPAPRPGQPAALGFTVALVWMSFCPLVAGAYIKLYCDLPRGDFEAWMDCAMV 541
DB 574 ASPLCLPAPRPGQPAALGFTVALVWMSFCPLVAGAYIKLYCDLPRGDFEAWMDCAMV 633
QY 542 RHVAMLTIFADGILCYPAVAFSPASMLGLPPVPEAVKSVLLVPLPACINPLLYLLEFP 601
DB 634 RHVAMLTIFADGILCYPAVAFSPASMLGLPPVPEAVKSVLLVPLPACINPLLYLLEFP 693
QY 602 HPRDDLRLRPRAAGSGPLAVALAAGELEKSSCDSTQALVAFSDVLLTLEASAGRPGLT 661
DB 694 HPRDDLRLRPRAAGSGPLAVALAAGELEKSSCDSTQALVAFSDVLLTLEASAGRPGLT 753
QY 662 TYGFPSTVLISQOQPARLEGSHCVERPENGFGNPQPSMDGELLRLRAGSTPAGGGLSG 721
DB 754 TYGFPSTVLISQOQPARLEGSHCVERPENGFGNPQPSMDGELLRLRAGSTPAGGGLSG 813
QY 722 GGGFQPSGLAFASHV 736
DB 814 GGGFQPSGLAFASHV 828

```

RESULT 2

```

ID LGR5_HUMAN STANDARD; PRT; 907 AA.
AC 075473; Q9UP75;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Leucine-rich repeat-containing G protein-coupled receptor 5 precursor
DE (orphan G protein-coupled receptor Hg38) (G protein-coupled receptor
DE 49).
GN GPR49 OR LGR5 OR GPR67.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98308104; PubMed=9642114;
RA McDonald T., Wang R., Bailey W., Xie G., Chen F., Caskey C.T.,
RA Liu Q.;
RT *Identification and cloning of an orphan G protein-coupled receptor of
RT the glycoprotein hormone receptor subfamily.";
RL Biochem. Biophys. Res. Commun. 247:266-270(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99065210; PubMed=9849958;
RA Hsu S.Y., Liang S.-G., Hsueh A.-J.W.;
RT Characterization of two LGR genes homologous to gonadotropin and
RT thyrotropin receptors with extracellular leucine-rich repeats and a G
RT protein-coupled, seven-transmembrane region.";
RL Mol. Endocrinol. 12:1830-1845(1998).
CC -!- FUNCTION: Orphan receptor. It may be an important receptor for
CC signals controlling growth and differentiation of specific
CC embryonic tissues (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Expressed in skeletal muscle, placenta, spinal
CC cord, and various region of brain.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -!- SIMILARITY: CONTAINS 17 LEUCINE-RICH REPEATS (LRR).

```

 CC This SWISS-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).

 CC EMBL: AF062006; AAC28019.1; -
 CC EMBL: AF061444; AAC77911.1; -
 CC Genbank: HGNC:4504; GPR49.
 CC MIM: 606667; -
 CC HSSP: P23945; 1XUN.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR00372; LRR_Nterm.
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR003591; LRR_typ.
 DR Pfam: PF00001; 7tm1; 1.
 DR Pfam: PF00560; LRR; 15.
 DR Pfam: PF01462; LRRNT; 1.
 DR PRINTS: PR00019; LEURICHRPT.
 DR SMART: SM00370; LRR; 3.
 DR SMART: SM00013; LRRNT; 1.
 DR PROSITE: PS00369; LRR_TYP; 8.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
 DR PROSITE: PS00262; G_PROTEIN_RECEP_F1_2; 1.
 DR PROSITE: PS00262; G_PROTEIN_RECEP_F1_2; 1.
 DR G-protein coupled receptor; Signal; Transmembrane; Glycoprotein;
 DR Repeat; Leucine-rich repeat.
 FT CHAIN 1 21
 FT SIGNAL 1 21
 FT DOMAIN 1 561
 FT TRANSMEM 562 582
 FT DOMAIN 583 593
 FT TRANSMEM 594 614
 FT DOMAIN 615 638
 FT TRANSMEM 639 659
 FT DOMAIN 660 682
 FT TRANSMEM 683 703
 FT DOMAIN 704 722
 FT TRANSMEM 723 743
 FT DOMAIN 744 767
 FT TRANSMEM 768 788
 FT DOMAIN 789 802
 FT TRANSMEM 803 823
 FT DOMAIN 824 907
 FT REPEAT 89 112
 FT REPEAT 114 136
 FT REPEAT 137 160
 FT REPEAT 162 184
 FT REPEAT 185 208
 FT REPEAT 209 232
 FT REPEAT 233 256
 FT REPEAT 257 279
 FT REPEAT 281 303
 FT REPEAT 304 327
 FT REPEAT 328 350
 FT REPEAT 351 375
 FT REPEAT 377 396
 FT REPEAT 397 420
 FT REPEAT 422 444
 FT REPEAT 444 464
 FT REPEAT 464 484
 FT REPEAT 484 504
 FT REPEAT 504 524
 FT REPEAT 524 544
 FT REPEAT 544 564
 FT CARBOHYD 63 63
 FT CARBOHYD 77 77
 FT CARBOHYD 208 208
 FT CARBOHYD 500 500
 FT CARBOHYD 592 592
 FT CARBOHYD 792 792
 FT CARBOHYD 90 90
 FT CONFLICT 212 212
 FT CONFLICT 212 212
 FT SEQUENCE 907 AA; 99997 MW; 822D5C5E6F0D9092 CRC64;

Query Match 47.9%; Score 1944; DB 1; Length 907;
 Best Local Similarity 52.2%; Pred. No. 5,6e-126;
 Matches 360; Conservative 113; Mismatches 186; Indels -30; Gaps 8;
 1 GHNLETLNANNKQIEFPVATRTGRLQELGFHNNNIKAIPKAFMGPILOTTHFYDN 60
 232 GLHSTLETLNANNKQIEFPVATRTGRLQELGFHNNNIKAIPKAFMGPILOTTHFYDN 291
 61 PLOFVGRSAFOYLPLKLTPLSLNGAMDIOEPDLKGTSTLEITLRTAGIRLPLSGMCOOL 120
 292 PLOFVGRSAFOYLPLKLTPLSLNGAMDIOEPDLKGTSTLEITLRTAGIRLPLSGMCOOL 351
 121 PRLVLELHNOIEELPSLRCKEIEIGLQHRIMEIGADTFSSQLQALDLSMAIR 180
 352 PNLQVLDLSYNLEDELPSEVSCOKLOKIDLRHNEIYEIKVDTEQQLSLSLWAMKIA 411
 181 SIHPEAFSTLHSLVKLDLPTNOLTTPLAGLGLIMHLKLGNALSGAFKDSFPKLRIL 240
 412 ITHNAFSTLPSLKLKLDLSNLLSPITGLHGLTHLKLGNHALQSLSENPELKI 471
 241 EEPYAYQCCPYGMCASFPEKASGQW-----EABDLHDEESSKRPGLGLARQAEHYDQ 294
 472 EMPYAYQCCAFYGCENAYKISNQMNKGDNSSMDLH-----KIDAGMPQADDE---R 520
 295 DLDELQLEM-EDSKPHRSVOCSPTRPGPKCEVLEESMGIRLAVMAIVLISVLCGLVL 353
 521 DLDELQLEM-EDSKPHRSVOCSPTRPGPKCEVLEESMGIRLAVMAIVLISVLCGLVL 580
 354 TYFAGGAPAPLPVYFVVGALAGANTLTGICGLASVADALTFQFSEYGARWETGLCRA 413
 581 TVFR-SPLYTSPILKILIGVLAANVMLTGVSAVLAVGADFTFGSEFARHAGWENGVC 639
 414 TGFVLAVGSEASVLLTLAVQCSVSCVRAVGSKPSLSVAGVIGCLALAGLAALP 473
 640 IGFSLFASSESVLTLALALERGFVSKAFETKAPFSSLLKILLCALLATLMAAVP 699
 474 LASVGEVAGSPICLPYAPBEGAPALGFTVALVMNSFCFLVAGAYIKLYCDLPBGDFE 533
 700 LLGSKRGASPLCLPL--PGEPSMGYMAVLIINSICFLMTLATYKLCNLDKQGLE 757
 534 AVMCAMRYHVAWMLIFADGLLYCPVAFSLFASMGLEFVTPBEAVKSVLLVPLPACINP 593
 758 NIMCQSNVKNHIALLLFNCLINCPVAFSLFSSLLNIFRISPEVIFILLVVPAPACINP 817
 594 LLYLLENPHEDDLRLRLPRA-----GDSGLAYAAAGLEKSSCDSDQALVAVSDVLI 648
 818 LLYLLENPHEDDLRLRLPRA-----GDSGLAYAAAGLEKSSCDSDQALVAVSDVLI 877
 649 LEASEAGRP---PGLFTYGFPSVTLISC 673
 878 YDLPSSVSPFAYPTESCHLSVAFFVC 906
 RESULT 3
 ID LGR5_MOUSE STANDARD; PRT; 907 AA.
 AC Q921P4; 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Leucine-rich repeat-containing G-protein-coupled receptor 5 precursor
 DE (G-protein-coupled receptor 49) (Orphan G-protein-coupled receptor
 DE FEY).
 GN GPR49 OR LGR5 OR FEY.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=99121227; PubMed=9920770;

RA Hermey G., Methner A., Schaller H.C., Hermans-Borgmeyer I.:
 RT Identification of a novel seven-transmembrane receptor with homology
 RT to glycoprotein receptors and its expression in the adult and
 RT developing mouse.".
 RT Biochem. Biophys. Res. Commun. 254:273-279(1999).
 CC -1- FUNCTION: Orphan receptor. It may be an important receptor for
 CC signals controlling growth and differentiation of specific
 CC embryonic tissues.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC and in the brain. In the central nervous system expression is
 CC restricted to the olfactory bulb. In the adrenal gland detected
 CC only in the neutral-crest derived chromaffin cells of the
 CC medulla, but not in the cells of the adrenal cortex. In the
 CC gonads, the expression is high in Graafian follicle, but absent
 CC from primary and secondary follicles.
 CC -1- DEVELOPMENTAL STAGE: Expressed from embryonic day 10.5 (E10.5) in
 CC the developing spinal cord and in the neuroepithelia of the
 CC myel-, met-, mes-, and diencephalon. Expression is transitory and
 CC the pattern changed rapidly.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -1- SIMILARITY: CONTAINS 17 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF110818; AAD14684.1; .
 DR HSSP: P23945; 1X0N.
 DR MGD: MG1:1341817; GPr49.
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_NTerm.
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR003591; LRR_1yp.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 15.
 DR PRINTS: PR001462; LRRNT; 1.
 DR PRINTS: PR00237; GPCR_RHODOPSIN.
 DR PRINTS: PR00019; LEURICHRPT.
 DR SMART: SM00370; LRR; 3.
 DR SMART: SM00013; LRRNT; 1.
 DR SMART: SM00369; LRR_TYP; 8.
 DR PROSITE: PS00237; G-PROTEIN_RECPT_FL_1; FALSE_NEG.
 DR PROSITE: PS0262; G-PROTEIN_RECPT_FL_2; 1.
 DR G-protein coupled receptor; Signal; Transmembrane; Glycoprotein;
 KW Repeat; Leucine-rich repeat.
 FT SIGNAL 1 21
 FT CHAIN 22 907
 FT DOMAIN 22 561
 FT TRANSSEM 562 582
 FT DOMAIN 583 593
 FT TRANSSEM 594 614
 FT DOMAIN 615 638
 FT TRANSSEM 639 659
 FT DOMAIN 660 682
 FT TRANSSEM 683 703
 FT DOMAIN 704 723
 FT TRANSSEM 724 744
 FT DOMAIN 745 767
 FT TRANSSEM 768 788
 FT DOMAIN 789 802
 FT TRANSSEM 803 823
 FT DOMAIN 824 907
 FT REPEAT 64 88
 FT REPEAT 89 112
 FT REPEAT 113 136
 FT REPEAT 137 160

FT	REPEAT	162	184	LRR 5.
FT	REPEAT	186	208	LRR 6.
FT	REPEAT	209	232	LRR 7.
FT	REPEAT	233	256	LRR 8.
FT	REPEAT	257	279	LRR 9.
FT	REPEAT	281	303	LRR 10.
FT	REPEAT	304	327	LRR 11.
FT	REPEAT	328	350	LRR 12.
FT	REPEAT	351	375	LRR 13.
FT	REPEAT	377	396	LRR 14.
FT	REPEAT	397	420	LRR 15.
FT	REPEAT	422	444	LRR 16.
FT	REPEAT	564	585	LRR 17.
FT	CARBOHYD	63	63	N-LINKED (GLCNAC...) (POTENTIAL).
FT	CARBOHYD	77	77	N-LINKED (GLCNAC...) (POTENTIAL).
FT	CARBOHYD	208	208	N-LINKED (GLCNAC...) (POTENTIAL).
FT	CARBOHYD	792	792	N-LINKED (GLCNAC...) (POTENTIAL).
FT	SEQUENCE	907 AA:	99681 MW:	5531676C0AAE253 CRC64;
SQ	SEQUENCE	907 AA:	99681 MW:	5531676C0AAE253 CRC64;

QY	1	GLHNETTLDLNNKIOEPVATRTIGLQIETGFHNNNKAIPKATMGNTLQTIHYDN	60
DB	232	GHLSLETDLNNNDLDEFFTAIKTSLNKLKELGPHNNKIPRAVGNPSLTTHFYDN	291
QY	61	PIQVGRSAFOYLPKILHTLSLNGAMDIOEPDLKGTTSLEILTRAGIRLPSGMOQL	120
DB	292	PIQVGRSAFOYLPKILHTLSLNGAMDIOEPDLKGTTSLEILTRAGIRLPSGMOQL	120
QY	121	PLRVLELSHNOIEFELPSLRHCOKLEIGLQNNRWEIGADTFPSQSLQALDSMNAIR	180
DB	352	PLRVLELSHNOIEFELPSLRHCOKLEIGLQNNRWEIGADTFPSQSLQALDSMNAIR	180
QY	191	SIHPEAFSLHSLVRLKIDLTDLNQTTLPLAGLGLMLKLGMLALSQAQSKDFPLRTI	240
DB	412	IIHPNAFSTLPKILKIDLSNLSLSPVYGLGHTLKLKGNRALSOLPSANFPELKII	471
QY	241	EVPRAYOCCPGMGKASFRKASQWEMEDLHDESSKRPGLLARQAEHNYDQDLEQ	300
DB	472	EMPSAYOCCAGGECENYKISNMNKKDDGNSVD-DLHKKDAGLFQVQDE----RDLEDFL	526
QY	301	LEM-EDSKRHPSPVOCSPPGPKCEYLFEESGIRLAWAVILSLVGLVLTFFAGG	359
DB	527	LDPEEDNALHSVQCSPPGPKCEYLFEESGIRLAWAVILSLVGLVLTFFAGG	359
QY	360	PAPLPYKFEVVGALAGANTLTGISCGLASVADLTFFQFSEYGAWEITGLCRATGFLAV	419
DB	586	PLYISIKILIGVIAVVDILMGVSSAVLAADAFPPGRAGAGAMWEDGICQIYGFSLI	645
QY	420	LGSEASVILTLAVVQCSVSVSCVRAVYKSPSLGSRVAGVAGCLADLAGLAALPLASVE	479
DB	646	FASSTIFLTLALEREFSVCSKFEVKAPLFSIRLIVLCVLLATLAIIPPLGSK	705
QY	480	YGASPLCLPYAPPEGOPALGFTVALVMNSFCFLVAGAYIKLYCDLPDGFEEVWCA	539
DB	706	YNASPLCLPL--PGGEPTSTGYWALVILNSLCELIIMATITKLYCSLEKGELEWMDOS	763
QY	540	MVRVAMITPFDGLICPVAFSLFASMLGDFPYVEAVKSVLLVLPPLACINPLLYLF	599
DB	764	MVKHIALILFANCLICPVAFSLFSSSLNLTFFISPDVIFILVLPPLSCINPLLYIF	823
QY	600	NPHRDLRLRLRPA-----GDSGLPAAAGELKSSCDSTQALVAVASVDLT--LEAS	652
DB	824	NPHKEDMGSKGTRKRWMSKASLISINSDDVEKSCSTQALVAVASVDLT--LEAS	883
QY	653	EAGRP--PGLETYGFSPVTLISC	673
DB	884	SGASPAVPMTESCHLSVAEVP	906
RESULT	4		

LGRA_RAT STANDARD: PRT: 951 AA.
 AC Q922H4: 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Leucine-rich repeat-containing G protein-coupled receptor 4 precursor.
 GN GPR48 OR LGRA4.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC NCBI_taxid=10116;
 RN SEQUENCE FROM N.A.
 RP TISSUE-Ovary;
 RX MEDLINE=99065210; PubMed=9849958;
 RA Hsu S.Y., Liang S.-g., Hsueh A.J.W.;
 RT "Characterization of two LGR genes homologous to gonadotropin and
 RT thyrotropin receptors with extracellular leucine-rich repeats and a G
 RT protein-coupled, seven-transmembrane region.";
 RL Mol. Endocrinol. 12:1830-1845(1998).
 CC -1- FUNCTION: Orphan receptor.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -1- SIMILARITY: CONTAINS 15 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC
 DR EMBL: AF061443: AAC77910.1: -;
 DR InterPro: IPR000216; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Nterm.
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR003591; LRR_tyr.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR_15.
 DR Pfam: PF01462; LRRNT; 1.
 DR PRINTS: PR00237; GPCR_Rhodopsn.
 DR PRINTS: PR00019; LEURICHRPT.
 DR SMART: SM00370; LRR; 4.
 DR SMART: SM00013; LRRNT; 1.
 DR SMART: SM00369; LRR_tyr; 5.
 DR PROSITE: PS000237; G-PROTEIN_RECP_FL_1; FALSE_NEG.
 DR PROSITE: PS00262; G-PROTEIN_RECP_FL_2; 1.
 DR G-protein coupled receptor; Transmembrane; Signal; Glycoprotein;
 KW Repeat; Leucine-rich repeat.
 FT CHAIN 1 24
 FT SIGNAL 1 24
 FT DOMAIN 25 544
 FT TRANSSEM 545 565
 FT TRANSSEM 566 575
 FT TRANSSEM 576 596
 FT TRANSSEM 597 619
 FT TRANSSEM 620 640
 FT TRANSSEM 641 661
 FT TRANSSEM 662 682
 FT TRANSSEM 683 703
 FT TRANSSEM 704 724
 FT TRANSSEM 725 756
 FT TRANSSEM 757 777
 FT TRANSSEM 778 783
 FT TRANSSEM 784 804
 FT TRANSSEM 805 951
 FT REPEAT 55 79
 FT REPEAT 81 103
 FT REPEAT 104 127
 FT LRR 1.
 FT LRR 2.
 FT LRR 3.

FT REPEAT 128 151 LRR 4.
 FT REPEAT 153 175 LRR 5.
 FT REPEAT 176 199 LRR 6.
 FT REPEAT 200 223 LRR 7.
 FT REPEAT 225 247 LRR 8.
 FT REPEAT 248 270 LRR 9.
 FT REPEAT 272 294 LRR 10.
 FT REPEAT 318 341 LRR 11.
 FT REPEAT 342 366 LRR 12.
 FT REPEAT 368 387 LRR 13.
 FT REPEAT 388 411 LRR 14.
 FT REPEAT 413 435 LRR 15.
 FT DISULFID 618 693 BY SIMILARITY.
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 505 505 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 951 AA; 104138 MW; EDD56AC072123461 CRC64;

Query Match 36.6%; Score 1410; DB 1; Length 951;

Best local Similarity 43.2%; Pred. No. 1.8e-94; Matches 303; Conservative 101; Mismatches 245; Indels 52; Gaps 5;

QY 1 GLNLETLDTNKLQEPFAIRTLGRLQELGFRNNIKAIPEKAFGNPLLOTIHFYDN 60
 DB 223 GLNLETLDTNLYNLYVEEFPQAIKALPSLKELGPHSNISVIYDAGFGNPLTIHLN 282
 QY 61 PLOFVGRSAFOYLPKLTLLSLNGAMDIOEFPDLKGTSLLETLTRAGIRLLPSGMCOOL 120
 DB 283 PLTFVGSAAHNLSDHLGLVIRGSLVQWEPNPLGVHLSLTLTGKISSIPDLCONQ 342
 QY 121 PRLVELSHNQIEELPSLRQCKLEBIGLQHNRIWIGADTESQLSSLOALDSNAIR 180
 DB 343 KMLRTLDLSYNNIRDLPSFGCRALIEISLRNQISLKEKTEGTLRLDLDSRLN 402
 QY 181 SIHPEARSTLSLVKLDLTNOLTLPLAGIGIMHKLGKGNLALSOAFSKDSEPKRLIL 240
 DB 403 EIHSGAFKAKLTITNLDVSNFELSPPTGELINGLNQIKLVGNFKALALANDRANRSL 462
 QY 241 EYPAVOCPEYMGCAFSPFKASGQWEADLHIDDESSKRPGLILARQENHYDODLDELQ 300
 DB 463 SVRYAYOCCAFWGCDSYANLNTEDNSPOEHVSFTEKGTADANVTSTAEENEHSQI 518
 QY 301 LEMEDSKPHSPVQCPPTGPPKPECEYLPESWGITLAWAYILSVLNGVLLTVFAGP 360
 DB 519 -----IIHCTPSVGAFFKPCPEYLLGSMWIRLITWFIPLVALLNLLVITVFA-SC 567
 QY 361 APLPEVYVGAIGANTLTIGISGLASVDAITFGOFSEYGAHWETGLGCRATGFLAVL 420
 DB 568 SLPASKFLFGLISVSNILMGITGITFLDAVSGRAEREGIMWETGSGCKVAGSLAVE 627
 QY 421 GSEASVLLTLAAVQCSVSVSCVRAVYKSPSLGVSRAVGLCLALAGIAALPLASGEY 480
 DB 628 SSESAVFLTLAAVERSVYFAKDLMKHGKSSHLRQFOVALALLLGAAYAGGFPFPHGOY 687
 QY 481 GASPLCLPYAPPEQAPALGFTVALVMNSFCPLVAGAYIKIVCDLRGDEAVADCM 540
 DB 688 SASPLCLPF--PTEPTPLGFTVTLNLSLAFLMALIYTKLVCKNLKEDISENSSQSV 745
 QY 541 VRHVAMLFADGLLYCPAFLSFASMLGLFVTEPAVKSVLVLYLPACINPLLYLLEN 600
 DB 746 IKHVAMLFTNCIFPCPAFLSFAPLITAISSPEIMKSVTLIFPPLACINPVLVYFEN 805
 QY 601 PHEFRDLRLRPRAG-----DSGPLAVYAAAGLEKSSCDSTQ 637
 DB 806 PKFKEDWMLKRRVTRKGSVSVSISOGGCEODEFYDDCGYSHLOGLTVYDCCESEFL 865
 QY 638 ALVAFSDVDTLIEASEAGRPGLTEYGFPSVTLISCOQGA 678
 DB 866 LTKPVSKHLI-----KSHSCPVLTAAACQRPBA 894

```

RESULT 5
LGR4_HUMAN
ID LGR4_HUMAN STANDARD; PRT; 951 AA.
AC Q9BX1; Q9NVD1;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Leucine-rich repeat-containing G protein-coupled receptor 4 precursor
   (G protein-coupled receptor 48).
GN GPR48 OR LGR4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=21294803; PubMed=11401528;
RA Loh E.D., Broussard S.R., Kolakowski L.F. Jr.;
RT "Molecular characterization of a novel glycoprotein hormone
   G-protein-coupled receptor."
RL Biochem. Biophys. Res. Commun. 282:757-764(2001).
CC -1- FUNCTION: Orphan receptor.
CC -1- TISSUE SPECIFICITY: Expressed in multiple steroidogenic tissues:
   placenta, ovary, testis and adrenal. Expressed also in spinal
   cord, thymoid, stomach, trachea, heart, pancreas, kidney, prostate
   and spleen.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -1- SIMILARITY: CONTAINS 15 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF346711; AAK31153.1;
DR EMBL: AF346709; AAK31153.1; JOINED.
DR EMBL: AF346710; AAK31153.1; JOINED.
DR EMBL: AF257182; AAF68989.1;
DR Genem: HGNC:13299; GPR48.
DR MIM: 606666;
DR InterPro: IPR00276; GPCR_Rhodpsn.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR003372; LRR_Nterm.
DR InterPro: IPR003592; LRR_Cyp.
DR InterPro: IPR003591; LRR_Cyp.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PF00560; LRR; 15.
DR Pfam: PF01462; LRRNT; 1.
DR PRINTS: PRO0237; GPCR_RHODOPSIN.
DR PRINTS: PRO0019; LEURICHRP.
DR SMART: SM00370; LRR; 6.
DR SMART: SM00013; LRRNT; 1.
DR SMART: SM00369; LRR_TYP; 15.
DR PROSITE: PS00237; G-PROTEIN_RECP_F1_1; FALSE_NEG.
DR PROSITE: PS50262; G-PROTEIN_RECP_F1_2; 1.
DR G-protein coupled receptor; signal; Transmembrane; Glycoprotein;
KW Repeat: Leucine-rich repeat.
FT SIGNAL 1 24
FT CHAIN 25 951
FT DOMAIN 25 544
FT TRANSSEM 545 565
FT DOMAIN 566 575
FT TRANSSEM 576 596
FT DOMAIN 597 620
FT TRANSSEM 621 641
FT DOMAIN 642 661

```

```

FT TRANSSEM 662 682
FT DOMAIN 683 703
FT TRANSSEM 704 724
FT DOMAIN 725 756
FT TRANSSEM 757 777
FT DOMAIN 778 783
FT TRANSSEM 784 804
FT DOMAIN 805 951
FT REPEAT 55 79
FT REPEAT 81 103
FT REPEAT 104 127
FT REPEAT 128 151
FT REPEAT 153 175
FT REPEAT 176 199
FT REPEAT 201 223
FT REPEAT 224 247
FT REPEAT 248 270
FT REPEAT 272 294
FT REPEAT 318 341
FT REPEAT 342 366
FT REPEAT 368 397
FT REPEAT 411 433
FT REPEAT 438 461
FT DISULFID 618 693
FT CARBOHYD 68 68
FT CARBOHYD 199 199
FT CARBOHYD 294 294
FT CARBOHYD 314 314
FT CARBOHYD 505 505
FT CONFLICT 292 292
FT CONFLICT 433 433
FT CONFLICT 668 668
SQ SEQUENCE 951 AA; 104460 MW; 5E0C2DFC22CA1BB CXC64;

```

```

Query Match 36.6%; Score 1409; DB 1; Length 951;
Best Local Similarity 42.7%; Pred. No. 2,1e-94;
Matches 300; Conservative 106; Mismatches 237; Indels 60; Gaps 9;

```

```

QY 1 GLHNETLDLVNKKLOEPFVATRTIGRLQELGFPHNNIKAIPEKAMGNPILQTHFYDN 60
DB 223 GLDNETLDLSTNNIGEPFOAIKARPSLKEIGFHSISIVIPDGADNPILRTTHLYDN 282
QY 61 PLQVGRSAFOYLPRKILHTLSLNGANDIOEPDLKGTSLIITLFRAGIRLLPSGACQOL 120
DB 283 PLSEVGNASAFHNLSDLSLVIGASWVOQFPNLGTGVHLESILITGTCKSSIPNNLCQD 342
QY 121 PLRLVLESHNOIEPLSLHRCQKLEBIGLQHRNWEIGADTFSSQLSLQALDLSWNAIR 180
DB 343 KMLRTLDLSTNNIRDLPSFNGCHALEEISLQRNOIYQIEGTFOGLISRLIDLSRNLH 402
QY 181 SIHPEAFSTLSLVKLDLTLDNQLTTLPLAGGGLMHKIKLMALISQANSKSPFKRLTL 240
DB 403 EIHSRAPATLPTINLDVSFNEILSFPEGLNLQNLKLVGFKEALAAFDVNLRL 462
QY 241 EYVYAYQCCPYGMCASFPEKASQWEADLHLDD---EESKRPLGLLARAQENHODL 296
DB 463 SVPYAYQCCAFWGCDSY-----ANLNTEDNSLQDHSVAQEKGRADANANTSTLENEHSDI 518
QY 297 DELQLEMDSKPHRSVQCSPTPGPKPCEYLFEESGIGLAWAVALVLSVLCGLVLLVF 356
DB 519 -----TIHCTPSTGAKEPCEYLLGSMIRLITWEIFLVALFENMLVILTF 564
QY 357 AGGPAPLPVYFVVGALAGANTLTIGISGLASVDALPFGQSEVGANMGLGCRANGF 416
DB 565 ASCSS-LPSSKTLFGLISVSNLPMGIYIGILTFLDVAVSWGRFAEGIMWETSGSGKVAGF 623
QY 417 LAVIGSASVILLTLAAVQCSVSCVAVAYKSPSLGSRVAGVLGCLALAGIAAALPLAS 476
DB 624 LAVFSSSAIFLLMLATAVERSLSAKDIKNGKSNHLKQFVALALLAFGLAVAGCFPLFH 683
QY 477 VGEFGASPLCLPYAPPPGQPAALGFVALVAMNSFCPLVAGATIKLYCDIPRGDFEAVW 536
DB 684 RGEYSASPLCLP--PYGEPISLGFVTVLNLSLAFLLMAVITYIKVCNDEKEDSENS 741

```

FT	TRANSEM	444	465	3 (POTENTIAL).
FT	DOMAIN	466	485	CYTOPLASMIC (POTENTIAL).
FT	TRANSEM	486	508	4 (POTENTIAL).
FT	DOMAIN	509	528	EXTRACELLULAR (POTENTIAL).
FT	TRANSEM	529	550	5 (POTENTIAL).
FT	DOMAIN	551	573	CYTOPLASMIC (POTENTIAL).
FT	TRANSEM	574	597	6 (POTENTIAL).
FT	DOMAIN	598	608	EXTRACELLULAR (POTENTIAL).
FT	TRANSEM	609	630	7 (POTENTIAL).
FT	DOMAIN	631	695	CYTOPLASMIC (POTENTIAL).
FT	REPEAT		68	LRR 1.
FT	REPEAT	69	93	LRR 2.
FT	REPEAT	119	143	LRR 3.
FT	REPEAT	170	192	LRR 4.
FT	REPEAT	193	216	LRR 5.
FT	REPEAT	218	240	LRR 6.
FT	DISULFID	442	517	BY SIMILARITY.
FT	CARBOHYD	191	191	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	199	199	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	293	293	N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE	695 AA;	78084 MW;	1EFDPFC046380D CXC64;
<hr/>				
Query Match 13.7%; Score 526.5; DB 1; Length 695;				
Best Local Similarity 26.1%; Pred. No. 1.5e-30;				
Matches 170; Conservative 108; Mismatches 290; Indels 83; Gaps 15;				

ID FSHR_SHEEP STANDARD; PRT; 695 AA.
 AC P35379; Q28573; Q28574; Q9TS19;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Follicle stimulating hormone receptor precursor (FSH-R) (Follicle stimulating hormone receptor).
 GN FSHR.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Caprinae; Ovis.
 OC NCBI_TaxID=9940;
 RX MEDLINE=93351750; PubMed=8394255;
 RA Yarney T.A., Sairam M.R., Khan H., Ravindranath N., Payne S., Seidah N.G.;
 RT "Molecular cloning and expression of the ovine testicular follicle stimulating hormone receptor.";
 RT Mol. Cell. Endocrinol. 93:219-226(1993).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM FSH-R1).
 RC TISSUE=Testis;
 RX MEDLINE=93351750; PubMed=8394255;
 RA Yarney T.A., Sairam M.R., Khan H., Ravindranath N., Payne S., Seidah N.G.;
 RT "Molecular cloning and expression of the ovine testicular follicle stimulating hormone receptor.";
 RT Mol. Cell. Endocrinol. 93:219-226(1993).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM FSH-R2), AND CHARACTERIZATION.
 RC STRAIN=Dorset-Leicester-Suffolk 1; TISSUE=Testis;
 RX MEDLINE=98031015; PubMed=9364440;
 RA Yarney T.A., Jiang L., Khan H., MacDonald E.A., Laird D.W.;
 RT "Molecular cloning, structure, and expression of a testicular follicle stimulating hormone receptor with selective alteration in the carboxy terminus that affects signaling function.";
 RT Mol. Reprod. Dev. 48:458-470(1997).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM FSH-R3), AND CHARACTERIZATION.
 RC STRAIN=Dorset-Leicester-Suffolk 1; TISSUE=Ovary;
 RX MEDLINE=20391225; PubMed=10527866;
 RA Babu P.S., Jiang L., Sairam A.M., Youyz R.M., Sairam M.R.;
 RT "Structural features and expression of an alternatively spliced growth factor type I receptor for follicle stimulating hormone in the developing ovary.";
 RT Mol. Cell Biol. Res. Commun. 2:21-27(1999).
 RN [5]
 RP SEQUENCE OF 1-51 FROM N.A.
 RX MEDLINE=98031017; PubMed=9364442;
 RA Sairam M.R., Subbarayan V.S.R.;
 RT "Characterization of the 5' flanking region and potential control elements of the ovine follicle stimulating hormone receptor gene.";
 RT Mol. Reprod. Dev. 48:480-487(1997).
 CC -1- FUNCTION: Receptor for follicle stimulating hormone. The activity of isoform FSH-R1 is mediated by G proteins which activate adenylate cyclase. Isoforms FSH-R2 and FSH-R3 also bind FSH, but this does not result in activation of adenylate cyclase. Isoform FSH-R3 may be involved in calcium signaling.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Plasma membrane (isoforms FSH-R1 and FSH-R2); Cell surface (isoform FSH-R3).
 CC -1- ALTERNATIVE PRODUCTS: 4 isoforms; FSH-R1 (shown here), FSH-R2, FSH-R3 and FSH-R4; are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: Isoform FSH-R3 is expressed in ovary and testis, but not in kidney.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC FSH/LSH/FSH SUPERFAMILY.
 CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L07302; AAA31525.1; -
 DR EMBL: L12766; AAA31523.1; -
 DR EMBL: L12767; AAA31524.1; -
 DR EMBL: L36115; AAK70667.1; -
 DR EMBL: AJ131735; CA01495.1; -
 DR EMBL: AF090438; AAC61749.1; -
 DR PIR: JG1493; JG1493.
 DR HSSP: P23945; 1XUN.
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Nterm.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 4.
 DR Pfam: PF01462; LRRNT; 1.
 DR SMART: SM00013; LRRNT; 1.
 DR PROSITE: PS00237; G_PROTEIN_RECPT_F1_1; 1.
 DR PROSITE: PS00262; G_PROTEIN_RECPT_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing.
 FT SIGNAL 1 17
 FT CHAIN 1 695
 FT DOMAIN 18 366
 FT TRANSMEM 367 387
 FT DOMAIN 388 398
 FT TRANSMEM 399 421
 FT DOMAIN 422 443
 FT TRANSMEM 444 465
 FT DOMAIN 466 485
 FT TRANSMEM 486 508
 FT DOMAIN 509 528
 FT TRANSMEM 529 550
 FT DOMAIN 551 573
 FT TRANSMEM 574 597
 FT DOMAIN 598 608
 FT TRANSMEM 609 630
 FT DOMAIN 631 695
 FT REPEAT 44 68
 FT REPEAT 69 93
 FT REPEAT 119 143
 FT REPEAT 170 192
 FT REPEAT 193 216
 FT REPEAT 218 240
 FT DISULFID 442 517
 FT CARBOHYD 191 191
 FT CARBOHYD 199 199
 FT CARBOHYD 293 293
 FT VARSPPLIC 126 133
 FT VARSPPLIC 135 695
 FT VARSPPLIC 224 259
 FT VARSPPLIC 260 695
 FT VARSPPLIC 643 670
 FT VARSPPLIC 671 695
 FT SEQUENCE 695 AA; 78237 MW; FBF75D89D8C0D4B CMC64;
 SQ
 QY 30 ELGFNNHNIKAIPEKAPMGNPLQTIHFYDNPV-QVFSRFAQVLPKLTSLNCGAMDIQ 88
 Db 50 ELRVLTKLRIVRPGARSGFDLKEIKLSQNDVLEIVANFSLPKLHEIRIEKANNLL 109
 Query Match 13.6%; Score 524.5; DB 1; Length 695;
 Best Local Similarity 26.4%; Pred. No. 2,le-30;
 Matches 175; Conservative 101; Mismatches 273; Indels 113; Gaps 18;

```

QY 89 EF-PDLKGTSLLEILTRAGIRLLPSCMCOOLPRLRVLELSHNOIEELPSLHRCQLEE 147
   || || || || || || || || || || || || || || || || || || || || ||
Db 110 YIDDP-----AFONLPRLKRLILSNITKIKHPRVHKIQSLQK 146
   || || || || || || || || || || || || || || || || || || || || ||
QY 148 I--GLQHN-RIMEIGADTFSSLS-SLQALDLSWNAIRSIHPEAF-----STL 190
   || || || || || || || || || || || || || || || || || || || || ||
Db 147 VLDDIODNINIHVTRNSFMGSLFESMIMVLSKNGIOETHNCAFNGTQDELNLSDNSNL 206
   || || || || || || || || || || || || || || || || || || || || ||
QY 191 HSL-----VKLLTDNQTLTLLAGLGLMLHLKLNLSQAFSKDSPFKLRI 239
   || || || || || || || || || || || || || || || || || || || || ||
Db 207 EELPNDVFOGASGVILLIDSTRIRSLSYGLNKLAKSTYHLKRLPSLEKVTYLV 266
   || || || || || || || || || || || || || || || || || || || || ||
QY 240 LEVPAYOCPPYGMCASEFKASGQWEADLA-----IDDEESSKRPGLLARQA 288
   || || || || || || || || || || || || || || || || || || || || ||
Db 267 ASLTPSPHOCAC-----ANWRQTSDLHPTCNKSLIRQEVDMWTQARGORISLAEDD 318
   || || || || || || || || || || || || || || || || || || || || ||
QY 289 ENHYODDIDDELQLEMEDS--KRPVSQCSPTGPPRPCEYLFEENQIRLAVNAIVL 346
   || || || || || || || || || || || || || || || || || || || || ||
Db 319 EPSYAKGPDMMYSEPDYDLCSEVNVVTCSPEDAFNCPCEDINGYDILRLVIMFTSLAIT 378
   || || || || || || || || || || || || || || || || || || || || ||
QY 347 CGLVLLTFVAGAPAPLPVKFVVGAIAGANTLTGIGSLASVDALTFGQSEYGARWE 406
   || || || || || || || || || || || || || || || || || || || || ||
Db 379 GNILVAV-ILITISQKLYPRFLMCLAFADLCIGYLLILLISVDVHTISQYHNVAIDMQ 437
   || || || || || || || || || || || || || || || || || || || || ||
QY 407 TGIGCRATGFLAVLGSSEASVLLLTAAVQ-----CSVSVSC--VRAYGKSPSLCSVRA 457
   || || || || || || || || || || || || || || || || || || || || ||
Db 438 TGAGCDAAEFVTFVASELSVYTLTALTIERMHTITHAMOLECKVHRAASITLVGVV-- 495
   || || || || || || || || || || || || || || || || || || || || ||
QY 458 GVLGCLALAGLAAALPLAVSGEASPLPLPVPREGOPALGETVALVMNSFCFLVVA 517
   || || || || || || || || || || || || || || || || || || || || ||
Db 496 -----FAFAVALPFIIFGISTYMKVSIPLM--DIDSPISQILVMSLVLANLAVVIC 546
   || || || || || || || || || || || || || || || || || || || || ||
QY 518 GAVIKLYCDLPGRDF-EAIVDCAMVRAVAMLLFPADGLLYCPVALFSPASMLGEPVTP 576
   || || || || || || || || || || || || || || || || || || || || ||
Db 547 GCYTHIYLVARNNTISSSDFTIAKRMAMLLFTDFLCMAPISEFPAISATKXPLITVSK 606
   || || || || || || || || || || || || || || || || || || || || ||
QY 577 VSVLLVLPPLACMLPLLYLFNPHPRDDLRKLRPRAGDSGLAATAAGELKSSCDST 636
   || || || || || || || || || || || || || || || || || || || || ||
Db 607 SKILLVFPYPCANPFLYAITFRNFRDFFIL-----LSKFGCYEV 649
   || || || || || || || || || || || || || || || || || || || || ||
QY 637 QA 638
   || || || || || || || || || || || || || || || || || || || || ||
Db 650 QA 651
   || || || || || || || || || || || || || || || || || || || || ||

```

```

RT *Porcine follicle-stimulating hormone receptor".
RL Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
CC -I- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC ADENYLATE CYCLASE.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC FSH/LSH/TSH SUBFAMILY.
CC -I- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L31966; AA086933.1; -.
DR EMBL: AF025377; AAC24981.1; -.
DR HSSP: P23945; 1XUN.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000372; LRR_Nterm.
DR Pfam: PR00001; 7tm.1; 1.
DR Pfam: PR00560; LRR; 2.
DR Pfam: PF01462; LRRNT; 1.
DR SMART: SM00013; LRRNT; 1.
DR PROSITE: PS00237; G-PROTEIN_RECP_F1_1; 1.
DR PROSITE: PS0262; G-PROTEIN_RECP_F1_2; 1.
DR G-protein coupled receptor: Transmembrane; Glycoprotein; signal;
DR Phosphorylation; Repeat; Leucine-rich repeat.
KW STGNAL
FT 1 17
FT CHAIN 18 695
FT DOMAIN 18 366
FT TRANSMEM 367 387
FT DOMAIN 388 398
FT TRANSMEM 399 421
FT DOMAIN 422 443
FT TRANSMEM 444 465
FT DOMAIN 466 485
FT TRANSMEM 486 508
FT DOMAIN 509 528
FT TRANSMEM 529 550
FT TRANSMEM 551 573
FT DOMAIN 574 597
FT TRANSMEM 598 608
FT DOMAIN 609 630
FT TRANSMEM 631 695
FT REPEAT 44 68
FT REPEAT 69 93
FT REPEAT 119 143
FT REPEAT 170 192
FT REPEAT 193 216
FT REPEAT 218 240
FT DISULFID 442 517
FT CARBOHYD 191 191
FT CARBOHYD 199 199
FT CARBOHYD 293 293
FT CONFLICT 2 2
FT CONFLICT 13 13
FT CONFLICT 60 60
FT CONFLICT 166 166
FT CONFLICT 215 215
FT CONFLICT 247 247
FT CONFLICT 257 257
FT CONFLICT 334 334
FT CONFLICT 349 349
FT CONFLICT 352 352
FT CONFLICT 383 383
FT CONFLICT 407 407
FT CONFLICT 421 421
FT CONFLICT 427 427

```

```

FT CONFLICT 435 435 D -> N (IN REF. 1).
FT CONFLICT 483 483 L -> V (IN REF. 1).
FT CONFLICT 550 550 T -> I (IN REF. 1).
FT CONFLICT 586 586 S -> V (IN REF. 1).
FT CONFLICT 607 607 A -> L (IN REF. 1).
FT CONFLICT 691 691 R -> H (IN REF. 1).
SQ SEQUENCE 695 AA; 78172 MW; E9EBDB29C79CA50 CRC64;

Query Match 13.6%; Score 523; DB 1; Length 695;
Best Local Similarity 26.1%; Pred. No. 2,7e-30;
Matches 171; Conservative 103; Mismatches 254; Indels 126; Gaps 17;

QY 30 ELGFHNNNKAIEKAPMGNDLQTHFYDNP1-QVGRSAFOYLPRKLTLSLNGAMDQ 88
DB 50 ELKFVLTAKVIRKAGFSGDLEKIRISQNDVLEIVANFSLPKLHEIRIEKANML 109
QY 89 EF-PDLKGTSLLEILTRAGIRLLPFGMCOQPLRVLLELSHNOIEELPLHCQKLE 147
DB 110 YIDPD-----AFQNLPLRLYLISNTGVKHLPAVHKIQSLQK 146
QY 148 I--GLOHN-RIMEIGADTFESQLS-SIQALDLSMNAIRSHPEARS-----TL 190
DB 147 VLLDIDQDNIHIVERNSEFVGLSEFESMILMSKNGIREIHNCARNGTQDELNDNDNL 206
QY 191 HSL-----VKLDLDNDQTLPLAGLGIMHLKLNALSOARSKDSFPKLR 239
DB 207 EELPNDVFGAGASPVLDLSRTIRHSLPSYGLNKLRLAKSTYMLKPLSLKEVTLME 266
QY 240 LEVPAVYQCCPYGMCASFPGASQOMEADLH-----LDDE 274
DB 267 ASLTYPSHCCAF-----ANMRQISDLHPICNKSILNQEVDVMTQAGQSVLAEDG 318
QY 275 ESKRRRLGLLAQAEHNYDODELEDMEDSKPHHSVQCSPPPGPKCEYLEESWGIR 334
DB 319 ESS-----LAEEFDMYSEFDYDLCNEVD-----VICSPEPDENPCEDIMGHDLR 366
QY 335 LAWMAVILLVSLNCGLVLTVEFGAPRLPVKEVGAIGAMTGLISGGLASVALT 394
DB 367 VLIWFISLAIIGN-IIVVILITTSQKYLTPRFKMCNLAFADLCIGIYLLASVDIHT 425
QY 395 FGQSEYCARMTGIGACRATGFLAVLGSSESVLLTLAA-----VOCSSVSC 442
DB 426 KIQYHNVALDMQTAGCDAGFEFTVASELSYVTLATFLERHHTITHAMQDCKV---- 481
QY 443 VRAYKPSLSGVBRGVGLCLAGLALPLASVGEYGSPLCLPRAPEGQPAIGFT 502
DB 482 -----QLRHAASIMLGMIFAFYVALPPIFGISSYMKVSLCLPM--DIDSPLSQLY 531
QY 503 VALVMNSFCPLVAVAGAYIKLYCDLPRCD-FAVWDCAMRVHAWLIFADGLYCPAEL 561
DB 532 VSLVLANVLAFAVVICGCVTHIYLRNPNIMSSSDTKIAKRAMLITFDLCMAPISEF 591
QY 562 SFASNLGLFPVTPPEVKSVLVPLPACLPDLVLLFNPHFEDLRRLPRAG 615
DB 592 AISAALKVPLITVSKILLVLFYPINSCANPELVAIFTKNFRDVEILLSKRG 645

RESULT 9
FSHR_HORSE STANDARD; PRT; 694 AA.
AC PA7799;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Follicle tropin receptor).
DE FSHR.
GN Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC TISSUE=Testis;
RX MEDLINE=94256960; PubMed=8198575;
RA Robert P., Amsellem S., Christophe S., Benifla J.L., Bellet D.,
RA Koman A., Bidaud J.M.;
RT "Cloning and sequencing of the equine testicular follitropin
RT receptor."
RL Biochem. Biophys. Res. Commun. 201:201-207(1994).
CC -!- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC ADENYLATE CYCLASE. AMONG ALL MAMMALIAN FSH RECEPTORS, ON THE HORSE
CC RECEPTOR DOES NOT BIND LH/CHORIONIC GONADOTROPIN (CG).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC FSH/LSH/TSR SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: S70150; AAB30854.1; -.
DR HSSP: P23945; 1XUN.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000372; LRR_Nterm.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PF00560; LRR; 3.
DR Pfam: PF01462; LRRNT; 1.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR SMART: SMO0013; LRRNT; 1.
DR PROSITE: PS00237; G-PROTEIN_RECPT_F1_1; 1.
DR PROSITE: PS50262; G-PROTEIN_RECPT_F1_2; 1.
DR KMW: G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
DR KMW: Phosphorylation; Repeat; Leucine-rich repeat.
DR KMW: SIGNAL.
FT CHAIN 1 18 694
FT DOMAIN 18 365
FT TRANSMEM 366 386
FT DOMAIN 387 397
FT TRANSMEM 398 420
FT DOMAIN 421 442
FT TRANSMEM 443 464
FT DOMAIN 465 484
FT TRANSMEM 485 507
FT DOMAIN 508 527
FT TRANSMEM 528 549
FT DOMAIN 550 572
FT TRANSMEM 573 596
FT DOMAIN 597 607
FT TRANSMEM 608 629
FT DOMAIN 630 694
FT REPEAT 44 68
FT REPEAT 69 93
FT REPEAT 119 143
FT REPEAT 168 192
FT REPEAT 193 216
FT REPEAT 218 240
FT DISULFID 441 516
FT CARBOHYD 191 191
FT CARBOHYD 199 199
FT CARBOHYD 268 268
FT CARBOHYD 293 293
SQ SEQUENCE 694 AA; 78004 MW; E2F077C5E8BCA54 CRC64;

Query Match 13.6%; Score 522; DB 1; Length 694;
Best Local Similarity 25.6%; Pred. No. 3,1e-30;
Matches 168; Conservative 113; Mismatches 282; Indels 92; Gaps 15;

QY 30 ELGFHNNNKAIEKAPMGNDLQTHFYDNP1-QVGRSAFOYLPRKLTLSLNGAMDQ 88

```



```

Db 50 ELRFVLTAKRVIPKGFSGFGLKEIEISQNDVLEVEIANVNSNLEPKLEIRERKANN-- 107
Oy 89 EPPDLKGTTSLELITRAGICILRLPSGMCQOOLPRVLLELSHNOIEELPSLRCKLEEI 148
Db 108 -----LXYIDHDAFQNLPNLOYLLISNGIKHLRAVHKIQSLQKV 147
Oy 149 --GLOHN-RIVEIGADTFQSLS-SIQALDLSMNAIRSIHPEAF-----STLH 191
Db 148 LLDIQDNINIHVERNSPGLSFESTIILRSKNGIOEINHCAFNQGLDELNTSYNNLE 207
Oy 192 SL-----VKLDLPDNOITLPLAGLGGLMLHLKLGKNAALSOAFSKDSFPKRL 240
Db 208 ELPNDFOGASCSPVILDISGTRIHSIPNGLNKKILRASTYNNLKKPLSEKFAVAMEA 267
Oy 241 EAVYAVOCCPGYGCASFKNAGS-----QWEAEDELHDEESSKRPUG 282
Db 268 NLTVPSHCFAF---AMRRROTSELOTTCKNSILROEYDMQARGERSLAEDDESSYPKG 324
Oy 283 LLARQAEHYDQDLDELQLEMEDSKRPHSVQCSPTPEKPECEYLFESKGIRLAVNAIVL 342
Db 325 FDMWYSFEED-----LCNEVVD-----VTCSPKPDAPNCEEDIMGYDLRLVLPFISI 373
Oy 343 LSVLCNGVLITVYFAGGAPRLPYKPYVGAAGANTLIGSCGLASVADALTFCQSFSEYG 402
Db 374 LAITGN-IIVLVILITTSQYKLYPRFLMCNLAFADLCIGYLLLLIASVDIHTKQYHNVA 432
Oy 403 ARMETGCGCRATGFLAVGSEASVLLTLAAVQCSVSVCVRAVYKSPSGSVRAVLGC 462
Db 433 IDMOTGAGCCDAGFTYFASLSLYTLATLERMHTTHAMOLECKYQLHNAASVMLVG 492
Oy 463 LALAGLAALPLASVGEYASPLCLPVPAPREGQRAALGFTVALVMNSFCFLVAVGATIK 522
Db 493 WIFAFVALLPIFISTYMKVSICLPM--DIDSPSLQILVMSLVLNLAIVYIGCVIL 550
Oy 523 LYCDLPGRDF-EAWDCMVRHVAWILFADGILYCPVAFSLPASMGLPVTPEAKSVL 581
Db 551 IYLVARNNIYSSSDMYIAKRMALIFTDFLCMAPISEFAISASLKVPLITVSKILL 610
Oy 582 LVVLPPLACNPLVLYLTFNPHRDDLRRLRPACDSGPLAVAAAELEKSSCDST 636
Db 611 VLEFYVINCAMPFLYAITKFRDFEILLKFG-----CYEMQOLYKTTST 660

RESULT 10
FSHR_MOUSE STANDARD: PRT: 692 AA.
AC P35378; Q9QWV8; Q9DAC2;
DT 01-JUN-1994 (Rel. 29, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Follicle
DE receptor).
DE FSHR.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv; TISSUE=Testis;
RA Teana-Sempere M., Manna P.R., Huhtaniemi I.T.;
RT "Molecular cloning of the mouse follicle stimulating hormone receptor
RT alternatively spliced variants and receptor inactivation by a C566T
RT transition in exon 7 of the coding sequence.";
RL Submitted (Sep-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Atakawa T., Hata A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,

```

b 676 RVTNSYVLVPL-NH 688

Protein	Phosphorylation; Repeat; Leucine-rich repeat.	Signal; Signal; cytochrome; signal;
W	1	POTENTIAL.
T	17	FOLLICLE STIMULATING HORMONE RECEPTOR.
T	18	EXTRACELLULAR (POTENTIAL).
T	18	1 (POTENTIAL).
T	365	CYTOPLASMIC (POTENTIAL).
T	387	2 (POTENTIAL).
T	397	EXTRACELLULAR (POTENTIAL).
T	420	3 (POTENTIAL).
T	421	CYTOPLASMIC (POTENTIAL).
T	442	4 (POTENTIAL).
T	443	EXTRACELLULAR (POTENTIAL).
T	464	3 (POTENTIAL).
T	465	CYTOPLASMIC (POTENTIAL).
T	484	4 (POTENTIAL).
T	507	EXTRACELLULAR (POTENTIAL).
T	508	4 (POTENTIAL).
T	527	EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 528 549 5 (POTENTIAL).
 FT DOMAIN 550 572 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 573 596 6 (POTENTIAL).
 FT DOMAIN 597 607 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 608 629 7 (POTENTIAL).
 FT DOMAIN 630 692 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 44 68 LRR 1.
 FT REPEAT 69 93 LRR 2.
 FT REPEAT 119 143 LRR 3.
 FT REPEAT 170 192 LRR 4.
 FT REPEAT 193 216 LRR 5.
 FT REPEAT 218 240 LRR 6.
 FT DISULFID 441 516 BY SIMILARITY.
 FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 692 AA; 77681 MW; 267EAT8C7CFD8C6 CRC64;

Query Match 13.3%; Score 513.5; DB 1; Length 692;
 Best Local Similarity 25.9%; Pred No. 1.3e-29;
 Matches 188; Conservative 106; Mismatches 284; Indels 147; Gaps 25;

DB 109 -----LYNPEAF-ONLPSRLILSLNTGKILPAVHKTSLOKV 147
 OY 149 --GLOHN-RIMEIGADYPSOLS-SLOALDSWNAIRSIHPEAF-----STLH 191
 DB 148 LLDQDNINIHIVARNSEFMGLSFESVILMLSKNGIEIHNCAGTOLDELNLSDNNLE 207
 OY 192 SL-----VKDLDTNQLTTLPLAGLGLMHLKLNLSQAFSKDSPKRL 240
 DB 208 ELPLVDFOGASGPIILDSRTKVSLSPLHGLNKLKLNARSTYRLKLPNDKVTTLMEA 267
 OY 241 EYPAAYOCCPRGMCASFEGKASGOWEADLH-----LDDEE 275
 DB 268 SLTPSHCAF-----ANLKRQISLHPICKNSILRODIDMTQIGDQVSLIDEP 319
 OY 276 SSKRPLGLIARQENHNDQIDLELOLEKEDSKPHPSVCCSPFGFKCYELFEWGLRL 335
 DB 320 S-----YKGSMDMYNEFDYDLGN--EYVD-----VTCSPKPDFAFNCEDIMGYNLRV 366
 OY 336 AVMALVILSVLCNGVILTLTFAGGPAPLPYKVGAIAGANTLTGISGLIASYDALTF 395
 DB 367 LWFSTILAITGNTLVV-VLTTSQYKLTVPFLMGNLAFADLCIGIYILLIASVDHTK 425
 OY 396 GOFSEYGARWETGLGCRATGFLAVLSEASVLLTLAVQ-----CSVSVSCVARYGKS 449
 DB 426 SQYHNYAIDWCGAGCDAAAGFTVFASELSTVYTLTAITLERHMTTHHAQLCC----KV 480
 OY 450 PSLGSVRAGVIGCLALAGLAALPLASVGEYASPLCLPYARPEGOPALGFTVALVMN 509
 DB 481 QLRHAASVWLD-WTPAPAAALPPIFGISSTYKVSICLPM--DIDSPLSQLYVMALLVN 537
 OY 510 SFCEPLVAGAVYIKLYCDLPGRDF-EAVWDCAMVRHVAWLIIPADGLIYCVAFISFPMIG 568
 DB 538 VLAFAVYICGCTHYITLVANPTIVSSSDTKIAKRMATLIFDFELCMADISFPAISASLK 597
 OY 569 LPPVTPPEAVKSVLLVLPAPCLNPLLYLFFNPHRBDRLRPRAGDSGPLVAAAGEL 628
 DB 598 VPLITVSKAKILLVETPNSCANPFLVAFIKNERRDFIL-----L 640
 OY 629 EKSSCDSTQALVAFSDVDLILEASEFARPPGLTETGFPSTLLISCOOPGARPLGSHCVC 668
 DB 641 SKFGCTEMQAOIYRT-----ETSSA-----THNF-HARKSHCS--SAPRTVNSVVLV 684
 OY 689 PEGNH 693
 DB 685 PL-NH 688

RESULT 12
 FSHR_MACFA ID FSHR_MACFA STANDARD; PRT; 695 AA.
 AC P32212;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE FSHR. Follicle stimulating hormone receptor precursor (FSH-R) (Follictrophin receptor).
 GN Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Macaca.
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA MEDLINE=94071854; PubMed=7504463;
 RX Gomoll J., Dankbar B., Sharma R.S., Nieschlag E.; "Molecular cloning of the testicular follicle stimulating hormone receptor of the non human primate Macaca fascicularis and identification of multiple transcripts in the testis."; Biochem. Biophys. Res. Commun. 196;1066-1072(1993).
 RL -1- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. FSH/LSH/FSH SUPERFAMILY.
 CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; X74454; CAA52463.1; .
 DR PIR; S36452; S36452.
 DR PIR; JN0898; JN0898.
 DR HSSP; P23945; 1XUN.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR00372; LRR_Nterm.
 DR Pfam; PF00001; 7tm_1; 1.
 DR Pfam; PF00560; LRR; 3.
 DR Pfam; PF01462; LRRNT; 1.
 DR SMART; SM00013; LRRNT; 1.
 DR PROSITE; PS00237; G_PROTEIN_RECPT_FL_1; 1.
 DR PROSITE; PS0262; G_PROTEIN_RECPT_FL_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; signal;
 KW Phosphorylation; Repeat; Leucine-rich repeat.
 KM
 FT SIGNAL 1 17
 FT CHAIN 18 695
 FT DOMAIN 18 366
 FT TRANSMEM 367 387
 FT DOMAIN 388 398
 FT TRANSMEM 399 421
 FT DOMAIN 422 443
 FT TRANSMEM 444 465
 FT DOMAIN 466 485
 FT TRANSMEM 486 508
 FT DOMAIN 509 528
 FT TRANSMEM 529 550
 FT DOMAIN 551 573
 FT TRANSMEM 574 597
 FT DOMAIN 598 608
 FT TRANSMEM 609 630

DR EMBL: S59900: AAB26480.1: -
 DR EMBL: M95489: AA52478.1: -
 DR EMBL: X68044: CA48179.1: -
 DR EMBL: S73199: AAB32071.1: -
 DR PIR: JN0122: JN0122.
 DR PDB: 1XUN: 15-MAY-97.
 DR Genew: HGNC: 3569: FSHR.
 DR MIM: 136435: -
 DR InterPro: IPR000276: GPCR_Rhodopsin.
 DR InterPro: IPR001611: LRR.
 DR InterPro: IPR00372: LRR_Nterm.
 DR Pfam: PF00001: 7tm_1: 1.
 DR Pfam: PF00560: LRR: 3.
 DR Pfam: PF01462: LRRNT: 1.
 DR SMART: SM00013: LRRNT: 1.
 DR PROSITE: PS00237: G_PROTEIN_RECPT_F1_1: 1.
 DR PROSITE: PS00262: G_PROTEIN_RECPT_F1_2: 1.
 DR G-protein coupled receptor: transmembrane; glycoprotein; signal;
 KW phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing;
 KW 3D-structure.
 KW SIGNAL: 1 17 POTENTIAL.
 FT CHAIN: 18 695 FOLLICLE STIMULATING HORMONE RECEPTOR.
 FT DOMAIN: 18 366 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM: 367 387 1 (POTENTIAL).
 FT TRANSMEM: 388 398 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM: 399 421 2 (POTENTIAL).
 FT DOMAIN: 422 443 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM: 444 465 3 (POTENTIAL).
 FT TRANSMEM: 466 485 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM: 486 508 4 (POTENTIAL).
 FT TRANSMEM: 509 528 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM: 529 550 5 (POTENTIAL).
 FT DOMAIN: 551 573 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM: 574 597 6 (POTENTIAL).
 FT TRANSMEM: 598 630 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM: 631 695 7 (POTENTIAL).
 FT DOMAIN: 696 93 CYTOPLASMIC (POTENTIAL).
 FT REPEAT: 119 143 LRR 1.
 FT REPEAT: 170 192 LRR 2.
 FT REPEAT: 193 216 LRR 3.
 FT REPEAT: 218 240 LRR 4.
 FT REPEAT: 241 260 LRR 5.
 FT DISULFID: 442 517 BY SIMILARITY.
 FT CARBOHYD: 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD: 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD: 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD: 318 318 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC: 224 285 MISSING (IN SHORT ISOFORM).
 FT VARSPLIC: 342 695 MISSING (IN SHORT ISOFORM).
 FT CONFLICT: 112 113 S -> R (IN REF. 4).
 FT CONFLICT: 112 113 N -> T (IN REF. 1).
 FT CONFLICT: 197 198 EL -> AV (IN REF. 1).
 FT CONFLICT: 295 295 S -> P (IN REF. 4).
 FT CONFLICT: 307 307 T -> A (IN REF. 1).
 FT CONFLICT: 680 680 N -> S (IN REF. 1).
 FT CONFLICT: 680 680 N -> S (IN REF. 1).
 FT SEQUENCE: 695 AA: 73BBE71F76D2C05 CMC64;
 SQ
 Query Match: 12.9%; Score 498.5; DB 1; Length 695;
 Best Local Similarity: 24.8%; Pred. No. 1.6e-28;
 Matches: 163; Conservative 116; Mismatches 284; Indels 93; Gaps 18;

QY 192 SL-----VKILDNDQTLTPLAGLGLMLHLKIKMLALSOAFSKDSKPPKRL 240
 DB 208 ELFNDVFGHSGVILDIIDIRLHSLPSYGLNKKLRARSTYNNKRLPTLEKVALMHA 267
 QY 241 EVDYAVQCCPYGMCASFPAASQGWAEDELH-----LDDESSKRPGLLARQAE 289
 DB 268 SLTPYSHCCAF-----ANMRQISELHPICNKLILQEDVDMQTGQSSLAEDNE 319
 QY 290 NHYDQDLDELQLEMEDSKPAP--SVQCSPTPGFPKCEYLFPESWIGRLAVNALVLSVLC 347
 DB 320 SSSRGDMTYTFEDVDLQNEVDVYDCSPKPDAFNPCEIDIMGYNILRVILWTFSLATLG 379
 QY 348 NGVLVLTVPAGPAPLPVYVGAIALAGANTLIGISGLASVDALTFQGFSEYGARMEI 407
 DB 380 N-ILVILVLTSTQYKLVPRFLMCLNAFADLCIGIYILLIASVDHITKQYHNYALDQCF 438
 QY 408 GLGRATGFLAVLGESEASVLLTLTAQV-----CSVSVCAVRAGKSPSGSVRAGVGL 461
 DB 439 GAGCDAAQFVTFVASELSVTLTATLERWHTITHAMQLDC-----KVQLRHASVYMWG 493
 QY 462 CLALAGLAAALPLASVGRYASPLCLPYAPPEQOPALGFVALVMNSCFVLVAGAYI 521
 DB 494 WI-PAFAALPPIGIGISSYMKVSTCLPA--DISPLSOLYMSILVNLVLAFAVIGCYI 550
 QY 522 KLYGDLPRGDF-EAWMDCAMVRHVMILFADGLLYCPVALSPASMLGLEPPYPEAVKSV 580
 DB 551 HYLTVRNPNIYSSSDIRIAKRAMMLFTDFLCMAPISFAISASIKVPLITVSKAKIL 610
 QY 581 LLVLPDLACINPLLYLFPNPFRODLRLPRAGDSPLAAYAAAGLEKSCST 636
 DB 611 LVLPFRINSCAPFLYAFITKFRDFEILLKCG-----CTEMQOLYKRIETST 661
 RESULT 14
 FSHR_EQUAS STANDARD: PRT: 687 AA.
 AC 095179;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Follicle stimulating hormone receptor precursor (FSH-R) (Follictrophin receptor).
 GN FSHR.
 OS Equus asinus (Donkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OC NCB1_TaxID=9793;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=97338913; PubMed=9195473;
 RA Richard F., Martinat N., Remy J.-J., Salesse R., Combarnous Y., *Cloning, sequencing and in vitro functional expression of a recombinant donkey follicle-stimulating hormone receptor: a new insight into the binding specificity of gonadotrophin receptors.*; RT Insignt into the binding specificity of gonadotrophin receptors.*; RL J. Mol. Endocrinol. 18:193-202(1997).
 CC - FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE CYCLASE.
 CC - SUBCELLULAR LOCATION: Integral membrane protein.
 CC - SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC - SIMILARITY: FSH/LSH SUBFAMILY
 CC - SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on way use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: U73659; AAB18245.1; -.

DR HSSP; P23945; 1XUN.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Nterm.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; 7tm_1; 1.
 DR Pfam: PF01462; LRRNT; 1.
 DR PRINTS: PR00237; GPCR_Rhodopsn.
 DR SMART: SM0013; LRRNT; 1.
 DR PROSITE: PS00237; G-PROTEIN_RECEP_F1_1; 1.
 DR PROSITE: PS00262; G-PROTEIN_RECEP_F1_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Phosphorylation; Repeat; Leucine-rich repeat.
 FT SIGNAL 1 17
 FT CHAIN 1 687
 FT DOMAIN 18 358
 FT TRANSMEM 359 379
 FT DOMAIN 380 390
 FT TRANSMEM 391 413
 FT DOMAIN 414 435
 FT TRANSMEM 436 457
 FT DOMAIN 458 477
 FT TRANSMEM 478 500
 FT DOMAIN 501 520
 FT TRANSMEM 521 542
 FT DOMAIN 543 565
 FT TRANSMEM 566 589
 FT DOMAIN 590 600
 FT TRANSMEM 601 622
 FT DOMAIN 623 687
 FT REPEAT 68 93
 FT REPEAT 93 119
 FT REPEAT 119 143
 FT REPEAT 143 170
 FT REPEAT 170 193
 FT REPEAT 193 216
 FT DISULFID 218 240
 FT CARBOHYD 434 509
 FT CARBOHYD 191 199
 FT CARBOHYD 199 293
 FT SEQUENCE 687 AA; 76937 MW; FC3AF0B5531DA9A CRC64;
 Query Match 12.88; Score 492.5; DB 1; Length 687;
 Best Local Similarity 25.18; Pred. No. 4.3e-28;
 Matches 167; Conservative 105; Mismatches 274; Indels 119; Gaps 19;
 30 ELGPHNNINAIPEKAFMGNPLLOTTFHYDNPY-QPYGRSAFOYLPKULHTLSUNGAMDIQ 88
 50 ELRFVLTKLRTVIRKGAFCSGFDLKKIEISONDVLEYEVANVFSNLPRLHEIRIEKANN-- 107
 89 EFPDLKGTSLLETLTLRAGIRLLPSPGMCQOLPRLRVLELSHNOIELEPSLHRCOKLEI 148
 108 -----LTYIDHDARQNPINLQYLLISNGIKHLPRAVHKIQSLQKV 147
 149 --GLQHN-RIMEIGADTFSGLS-SLGADLDSWMAIRSHPEAF-----STLH 191
 148 LLDIDQNMIIHYVERNSFNGSLTSLRSLKNGIOEHNCAFNGTOLDLNLSDNNIE 207
 192 SJ-----VKLDLTNOJLTPLAGLGLMLKKNLALSOQSKSPFKRL 240
 208 ELRPNDVFGASGPVILDISGTRISHLPNGYLENKKLRARSTYNNLKKPLSEKFAALMDA 267
 241 EEPYVAYOCPCPGMCAFPKASQWMAEDLHD-----DESSKRPLGLARQA 288
 268 SLTYSHSC-----AFANMRQQTSELQTCNKSLIRQEVMTQARGERSVLAEDD 317
 289 ENHYDQDL-ELQLEMEDSKPHSPVOCPTPGPKPCYELFESWQIRLAVAVIALLSYLC 347
 318 ESMWSEFDYDLCNEVVD-----VTCSPKPDAPNFCEDINGYDILRYLIMFISILATG 371
 348 NGVILLITVFAGGAPRLPVPYKVFVAGAIAGANTLTGISGLASVALTGQFSEYARRET 407
 372 N-IIVLVLLITSQYKLTVPRLMCMNLAFAADLCIGIYLLIASVDIHTKSQYHNVAIDMOT 430

QY 408 GLGCRATGFLAVLGSEASVLLTLAA-----VCCSVSY---SCVRAVGKSPSL 452
 DB 431 GAGCDAAAGFFTVFEGSELSTVYTLTAITLERMHTINHAQLEKQVLLRRASAMLVGMITGE 490
 QY 453 GSVRAGVIGLALAGLAALPLASVGEYASPLCLPYAPPEGOPALGFTVALVMNMSFC 512
 DB 491 G-----VGL-----LPFGISTYMKVSVSLCPM--DIDSPLSOLYMSLIVLVLA 533
 QY 513 FLVYAGAYIKLYCDUPRGDF-EAWMDCANVRHVAWLIFADGLLYCPYAFSLFASMTGLFP 571
 DB 534 FVVICGGYTHIYLVTRNPNIVSSSDPKIKAKRMGILITFTPLCMAPLSFEGISASLVAL 593
 QY 572 VPEEAVKSVLLVPLPACINPLLLYLFNHFDDLRRLRPRAGDSGPLAYAAAGELEKS 631
 DB 594 IIVSKSKILVLELYPNSCANPLVAILFKNFRDFFILSKFG-----CYEQAOQTYRT 648
 QY 632 SCDDST 636
 DB 649 EHSST 653

RESULT 15
 TSHR_BOVIN
 ID TSHR_BOVIN STANDARD; PRT; 763 AA.
 AC Q27987;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Thyrotropin receptor precursor (TSH-R) (Thyroid stimulating hormone receptor).
 GN TSHR.
 OS Bos taurus (Bovine).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC OX NCBI_Taxid=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thyroid;
 RX MEDLINE=97280128; Pubmed=9134497;
 RA Silverides D.W., Houde A., Ethier J.F., Lussier J.G.;
 RT "Bovine thyrotropin receptor cDNA is characterized by full-length and truncated transcripts."
 RL J. Mol. Endocrinol. 18:101-112(1997).
 CC - FUNCTION: RECEPTOR FOR THYROTROPIN. PLAYS A CENTRAL ROLE IN CONTROLLING THYROID CELL METABOLISM. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE CYCLASE.
 CC - SUBCELLULAR LOCATION: Integral membrane protein.
 CC - SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC FSH/LSH/TSR SUBFAMILY.
 CC - SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC OR send an email to license@isb-sib.ch.
 DR EMBL: U15570; AAC18639.1; -.
 DR HSSP: P16473; 1XUM.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 1.
 DR PROSITE: PS00237; G-PROTEIN_RECEP_F1_1; 1.
 DR PROSITE: PS00262; G-PROTEIN_RECEP_F1_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Phosphorylation; Repeat; Leucine-rich repeat.
 FT SIGNAL 1 21
 FT POTENTIAL.

```
FT CHAIN 22 763 THYROTROPIN RECEPTOR.
FT DOMAIN 22 412 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 413 440 1 (POTENTIAL).
FT DOMAIN 441 449 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 450 472 2 (POTENTIAL).
FT DOMAIN 473 493 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 494 516 3 (POTENTIAL).
FT DOMAIN 517 536 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 537 559 4 (POTENTIAL).
FT DOMAIN 560 579 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 580 601 5 (POTENTIAL).
FT DOMAIN 602 624 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 625 648 6 (POTENTIAL).
FT DOMAIN 649 659 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 660 681 7 (POTENTIAL).
FT DOMAIN 682 763 CYTOPLASMIC (POTENTIAL).
FT REPEAT 51 74 LRR 1.
FT REPEAT 125 150 LRR 2.
FT REPEAT 151 174 LRR 3.
FT REPEAT 176 199 LRR 4.
FT REPEAT 201 223 LRR 5.
FT REPEAT 225 248 LRR 6.
FT DISULFID 493 568 BY SIMILARITY.
FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 763 AA; 86431 MW; 35E9F647E7ED7A8C CRC64;
```

Query Match 12.0%; Score 463; DB 1; Length 763;

Best Local Similarity 24.1%; Pred. No. 6.7e-26; Matches 173; Conservative 118; Mismatches 310; Indels 116; Gaps 19;

```
QY 37 NIKAIPEKAFNGNPLQTIHEYDNPIDFVGSAFOYLPKLTLSLNGAMDIOE----PP 91
   ::::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 43 DIQSIPSLP-----PSTQTLKFIEHLTKTIPSAFSLNPISRIYLSIDATLQQLSHSFY 98
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 92 DLKGTTSLEILTTRAGIRLPSCMCQOLPRLVLELSHNQIEELPSLHR----- 141
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 99 NLSKVTHIEIRN-TRS-LTYIDSGALKELPLKFLGIFNTGLRVFPDLTKIYSTVDFITL 156
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 142 -----COKLEIIGLOHNRIWEIGADTFQSLSLQALDLSMNA-IRS 181
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 157 EITNPNPYWTSIPANAFOGLCNETTLTKLYNNGFTSIQGHAFNG-TKLDAYLKNKRYLV 215
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 182 IHPFAFTIHS-LVKLDLTDNQLTTLPLAGLGIMHLKLGMLALSOAFSKDSFPKRLIL 240
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 216 IGQDAFAGVYSGPTLLDISYTSVTALPSKGLEHLKELLARNWTLPKLPLSLFLHLTRA 275
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 241 EYPAIYQCCPYG-----MC-----ASFFKASGQWEADLHLDDEE 275
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 276 DLSYPSHCACARKNOKKIRIGLQSLMCNESSIRGLRORKSASALNGPFQYEDLIG-DGSA 334
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 276 SSKRPLGLLAROAHNY-----DQDDEL-----QLEMEDSKRHPSY----- 312
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 335 GYKENSKEQDTQSNSHYVFEEDDEIIGEQQLKNPOEFTLQAFDSHYDYTVCGGSED 394
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 313 -QCSPTPGPFKPCXYLFESWGIRLAWAIVLLSVLCNGVLTLTFVAGGAPALPPVKFVVG 371
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 395 MVTCPKSDENPCEDIMYKFLRIYVWFVSLALGNGVFLVILLTSHYKLTVP-RFLMC 453
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 372 AIAQANTLTGISCGLLASVDALTGQFSEYGARWETGLGCRATGFLAVLGSEASVLLTL 431
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 454 NLAFADFCMGYLLILLASVDLYTOGEYVNHAIQMOTGPCNTAGFTVPASELSVYTLTV 513
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 432 AAQCSVSVSCVRAVGRKPSLGSVRAVLGCLALAGLAALPLASGEYGAAPLCPLPYAP 491
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 514 ITLERMHAITPAMRLDRKIRIMHAYVIMGGWCCFLALLPLVGLISSYAKVSICLPM-- 571
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 492 PEGOPALGFTVALVVMNMSFCFLVAVAGAYIKLYCDLPR-----GDFEAVWDCAMVRHVA 546
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 572 DFTPLALAYITIVLLINIAFIIVACACVYKITYIVRNPHYNGD-----KDTRIAKRMAY 627
```

```
QY 547 LIFADGLICPVAFLEASAMGLFPPVPEAVKSVLLVYLPACLNPLLYLLENPHFRD 606
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 628 LIFDFECMAPISTFYALSALMNNKPLITYTNSKILVLFFYPLNSCANPFLAIFTAFORD 687
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 607 LRRLRPRAGDSGPLAYAAAGE-----LEKSSCDSTQATVAFSVDVDLLEAS 652
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 688 VFMLLSKFGICKRQAQAVRGGRVSPKNSGTGLRVQKVPPDVQSLPNVQDDYELLENS 744
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: November 8, 2002, 19:34:59
Job time : 16.7524 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2002, 19:32:55 ; Search time 18.8167 Seconds
(without alignments)
3760.229 Million cell updates/sec

Title: US-09-851-595-8
Perfect score: 3850
Sequence: 1 GLHNLETLDTLNTYKQLQEPV.....GGLSGGGGFPQSLAFASHV 736

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1844	47.9	907	2	JE0176	orphan G protein-c
2	1759	45.7	907	2	JG0193	foliitropin recept
3	534	13.9	694	2	JC4301	follicle stimulat
4	526.5	13.7	695	2	I45896	foliitropin recept
5	524.5	13.6	695	2	JC1493	foliitropin recept
6	522	13.6	694	2	JC2237	foliitropin recept
7	513.5	13.3	692	2	A34548	foliitropin recept
8	502.5	13.1	695	1	JN0898	foliitropin recept
9	498.5	12.9	695	1	ORHUP	foliitropin recept
10	463	12.0	696	2	JC7361	thyrotropin recept
11	460	11.9	764	2	A40077	thyrotropin recept
12	456.5	11.9	696	2	A41344	thyroid stimulat
13	456.5	11.9	764	2	JC5643	thyrotropin recept
14	454.5	11.8	764	2	I48882	thyrotropin recept
15	449.5	11.7	700	2	A42395	lutropin receptor
16	443.5	11.5	700	2	I77463	lutropin-choriogon
17	443.5	11.5	700	2	A49744	thyroid stimulat
18	442.5	11.5	793	3	JC7390	thyrotropin recept
19	441	11.5	764	1	ORHUP	thyroid stimulat
20	432.5	11.2	814	2	JC7389	thyrotropin recept
21	432.5	11.2	889	2	T20123	hypothetical prote
22	424.5	11.0	764	2	A35956	thyrotropin recept
23	413	10.7	699	1	ORHUP	lutropin-choriogon
24	384	10.0	925	1	JC2037	G protein coupled
25	338	8.8	1115	2	S40241	glial cell membran
26	248	6.4	1091	2	A58532	hypothetical prote
27	240.5	6.2	1066	2	T15864	hypothetical prote
28	239.5	6.2	738	2	T19938	hypothetical prote
29	237	6.2	653	2	T25194	hypothetical prote

30	233.5	6.1	536	2	A34901	lysine carboxypept
31	231.5	6.0	1385	2	T13887	tlr protein - trit
32	230.5	6.0	1389	2	T13852	gene wheeler prote
33	227.5	5.9	603	2	JC6128	insulin-like growt
34	226	5.9	605	2	JC5239	insulin-like growt
35	226	5.9	605	2	A41915	insulin-like growt
36	224.5	5.8	575	2	T29972	hypothetical prote
37	222.5	5.8	575	2	JC1282	insulin-like growt
38	222	5.8	603	2	AD1822	leucine-rich repea
39	220	5.7	1039	2	T22117	hypothetical prote
40	214	5.6	1531	2	T42218	slit-1 protein hom
41	213	5.5	559	2	T42998	Ras-binding protei
42	206.5	5.4	572	2	T30947	hypothetical prote
43	206.5	5.4	613	2	A88684	protein ACT 2 [imp
44	206	5.4	1523	2	T13953	MEFs protein - ra
45	203.5	5.3	717	2	T33295	hypothetical prote

ALIGNMENTS

RESULT 1

JE0176 Orphan G protein-coupled receptor precursor - human

C:Species: Homo sapiens (man) C:Date: 03-Jul-1998 #sequence-revision 10-Jul-1998 #text-change 21-Jul-2000

C:Accession: JE0176

R:McDonald, T.; Wang, R.; Bailey, W.; Xie, G.; Chen, F.; Caskey, C.T.; Liu, Q. Biochem. Biophys. Res. Commun. 247, 266-270, 1998

A:Title: Identification and cloning of an orphan G protein-coupled receptor of the gl

A:Reference number: JE0176; MID:98308104; PMID:9642114

A:Accession: JE0176

A:Molecule type: mRNA

A:Residues: 1-907 <MCD>

A:Cross-references: GB:AF062006; MID:93366801; PIDN:AAC28019.1; PID:93366802

C:Comment: This protein is a receptor for a novel class of glycoprotein ligands.

C:Genetics:

A:Gene: H638

F:1-21/Domain: signal sequence

F:562-583/Domain: transmembrane #status predicted <TM1>

F:594-616/Domain: transmembrane #status predicted <TM2>

F:639-660/Domain: transmembrane #status predicted <TM3>

F:681-701/Domain: transmembrane #status predicted <TM4>

F:725-744/Domain: transmembrane #status predicted <TM5>

F:768-791/Domain: transmembrane #status predicted <TM6>

F:803-824/Domain: transmembrane #status predicted <TM7>

Query Match 47.9%; Score 1844; DB 2; Length 907;
Best local similarity 52.2%; Pred. No. 3,1e-133;
Matches 360; Conservative 113; Mismatches 186; Indels 30; Gaps 8;

QY	1	GLHNLETLDTLNTYKQLQEPVATIRTLGRLOELGFHNNNKAIPKAFMKNPILQTHFNDN	60
DB	232	GLHSLETLDTLNNINDEPPTAIRLSNKEGFHNNNKAIPKAFMKNPILQTHFNDN	291
QY	61	PIQFGRAFOYLPKLTLSUNGAMDIOEPDKGTSELTITRAGIRLPSGMCQOL	120
DB	232	PIQFGRAFOYLPKLTLSUNGAMDIOEPDKGTSELTITRAGIRLPSGMCQOL	351
QY	121	PLRLVLESHNOIEELPSLRHCOKLEIGLQHNRIWEIGADFFSOLSSIQALDLSNNAIR	180
DB	352	PLRLVLESHNOIEELPSLRHCOKLEIGLQHNRIWEIGADFFSOLSSIQALDLSNNAIR	411
QY	181	SIHPEAFSTHSLVKIDLTNDQTLTPLAGLGLMHTKIKGNLALSQAFKSPKRLIT	240
DB	412	SIHPEAFSTHSLVKIDLTNDQTLTPLAGLGLMHTKIKGNLALSQAFKSPKRLIT	471
QY	241	EVPYAVGCCPYGMCASFASGQW-----EADLHLDDESSKRPGLILARQENHYDQ	294
DB	472	EVPYAVGCCPYGMCASFASGQW-----EADLHLDDESSKRPGLILARQENHYDQ	520
QY	295	DLDELQLEM-EDSKRPHSVQSCSPFGPKPCYELFESWGIRLAWVAIVLLSYLNCGLVLL	353

```

Db 521 DLEDFELDFEEDLKALHSVOCSPSPGPFKPECHLLDGLWLTIGWTFIVALTALTCALVTS 580
OY 354 TVFAGGPAFLPYKVFYVAGTAGANTLTIGISGLLASVDALTFGOFSEYGANWETGLGCR 413
Db 581 TVFR-SPLYSPLTIGVIAVMMMLTGVSSAVLAGVDAFTFGSPARHAGWENGVGHV 639
OY 414 TGFVLAVGSEAVLLTLTAVVOCSSVSCVAVYKSPSLGSRVAGVGLCLAGLAAALP 473
Db 640 IGFSLIFASSESVFLLTLALERGFYSVYKAKFTKAPFSSLKYLITLICALLALMAAVP 699
OY 474 LMSVGEASPLLPAPRPGOPALGFTVALVMNSFCFLVAVAGYIKLYCDLPRGDFE 533
Db 700 LUGSKYKASPLCLPL--PGESEPTMGVWALITLNSLCFLMTIATYTKLYCNDKGDLE 757
OY 534 AVWDCAVHRVAVMLIFADGLYCPVAFSLFASMSLGLFPYPAVYKSVLVAVLPACLANP 593
Db 758 NIMDCSMVKHIALLTFTNCLNCPVAFSLFSSSLINTFTISEVYIKFTLLVVPPLACLANP 817
OY 594 LLYLFNPFRRDDLRLRPPRA-----GDSGPLAVYAAAGELEKSSCDSTQALVAFSDVLI 648
Db 818 LLYLFNPFRRDKEDLVSLRKQTYVWTRSKHPSLMSINSDDVEKQSCDSTQALVFTFSSSIT 877
OY 649 LEASEGRP-----PGLFTYGPSPVTLISC 673
Db - 878 YDLPSSVSPAPVAPVPTESCHLSVAFAVPC 906

```

RESULT 2

```

G:protein-coupled receptor FFX - mouse
C:Species: Mus musculus (house mouse)
C>Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000
C:Accession: JG0193
R:Hermey, G.; Methner, A.; Schaller, H.C.; Hermans-Borgmeyer, I.
Biochem. Biophys. Res. Commun. 254, 273-279, 1999
A>Title: Identification of a novel seven-transmembrane receptor with homology to glycophorin A
A:Reference number: JG0193; MUID:99121227; PMID:9920770
A:Accession: JG0193
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-907 <HER>

```

Query Match 45.7%; Score 1759; DB 2; Length 907;

Best Local Similarity 51.1%; Pred. No. 1e-126; Indels 18; Gaps 8;

```

Matches 349; Conservative 121; Mismatches 195; Indels 18; Gaps 8;
OY 1 GLHLETLIDLNVNKLQEPVPAIRTLQELGFHNNNIKAIFEKAFMGNPLQTHFYDN 60
Db 232 GLHLETLIDLNVNKLQEPVPAIRTLQELGFHNNNIKAIFEKAFMGNPLQTHFYDN 291
OY 61 PIGVGRSAFOYLKRLHTLSNGAMDIOEPDLKGTSLSEITLTRAGIRLLPSGMCQOL 120
Db 292 PIGVGRSAFOYLKRLHTLSNGAMDIOEPDLKGTSLSEITLTRAGIRLLPSGMCQOL 351
OY 121 PRLVLELSHNQIELDLSLRQCKLEIGIQHNRIMETIGADTFSSQLQALDLSWNAIR 180
Db 352 PRLVLELSHNQIELDLSLRQCKLEIGIQHNRIMETIGADTFSSQLQALDLSWNAIR 411
OY 181 SIHPEASTHSLVKLTLNQTLTLPLAGLGGIMHLKLGKALISQAQSKDPSFKRLTL 240
Db 412 SIHPEASTHSLVKLTLNQTLTLPLAGLGGIMHLKLGKALISQAQSKDPSFKRLTL 471
OY 241 EYVYAVQCCPYGMCASFPAKSGQWEAEDLHDDDESSKRRPGLLARQENHYDQDLEQ 300
Db 472 EYVYAVQCCPYGMCASFPAKSGQWEAEDLHDDDESSKRRPGLLARQENHYDQDLEQ 526
OY 301 LEM-EDSKPSPVQCCSTPPGPFKPECEYLFESWGIRLAVMAIVLLSVLNGVLLTFVAGG 359
Db 527 LDEEEDLNALHSVQCSPPGPFKPECEYLFESWGIRLAVMAIVLLSVLNGVLLTFVAGG 585
OY 360 PADLPYKFFVAGIAGANTLTIGISGLLASVDALTFGOFSEYGANWETGLGCRATGFLAV 419

```

```

Db 586 PLYISSIKLLIGVIAVDILMGVSAVLAADAFTFGFAOHGAMWEDGIGCQIVFLSI 645
OY 420 LGSSEAVLLTLTAAVOCSSVSCVRAVAKSPISGVSAGVGLGCLAGLAAALPLASVGE 479
Db 646 PASESSIFLTLTAAERGSVSKSSKFEVAPLPSLAIYLLCVLTLTATTPLLGSK 705
OY 480 YGASPLCLPYAPREGOPALGFTVALVMNSFCFLVAVAGYIKLYCDLPRGDEAVMDCA 539
Db 706 YGASPLCLPYAPREGOPALGFTVALVMNSFCFLVAVAGYIKLYCDLPRGDEAVMDCA 763
OY 540 MYRNVAMLEFADGLYCPVAFSLFASMSLGLFPYPAVYKSVLVAVLPACLANP 599
Db 764 MYRNVAMLEFADGLYCPVAFSLFASMSLGLFPYPAVYKSVLVAVLPACLANP 823
OY 600 NPHFRDDLRLRPPRA-----GDSGPLAVYAAAGELEKSSCDSTQALVAFSDVLI--LEAS 652
Db 824 NPHFRDDLRLRPPRA-----GDSGPLAVYAAAGELEKSSCDSTQALVAFSDVLI--LEAS 883
OY 653 EAGRP--PGLFTYGPSPVTLISC 673
Db 884 EAGRP--PGLFTYGPSPVTLISC 906

```

RESULT 3

```

JG4301
N:Alternates names: follicle-stimulating hormone receptor
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 16-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 21-Jan-2000
C:Accession: JG4301
R:Remy, J.J.; Labib-Mansais, Y.; Yerle, M.; Bozon, V.; Couture, L.; Pajot, E.; Grebe
Gene 163, 257-261, 1995
A>Title: The porcine follicle-stimulating hormone receptor: cDNA cloning, functional expression and characterization
A:Reference number: JG4301; MUID:96011644; PMID:7590277
A:Accession: JG4301
A:Molecule type: mRNA
A:Residues: 1-694 <REMA>
A:Cross-references: GB:L31966
A:Experimental source: ovary granulosa cells
C:Comment: This receptor belongs to the family of the G-protein coupled receptors. It
C:Genetics:
A:Gene: fshr
A:Map position: 3 q2.2-q2.3
C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repa
C:Keywords: G protein-coupled receptor; hormone receptor; transmembrane protein
F:1-365/Domain: follicle-stimulating hormone binding #status predicted <HOB>
F:366-388/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>
F:398-420/Domain: transmembrane #status predicted <TM1>
F:443-464/Domain: transmembrane #status predicted <TM2>
F:485-507/Domain: transmembrane #status predicted <TM3>
F:528-549/Domain: transmembrane #status predicted <TM4>
F:573-596/Domain: transmembrane #status predicted <TM5>
F:608-629/Domain: transmembrane #status predicted <TM6>

```

Query Match 13.9%; Score 534; DB 2; Length 694;

Best Local Similarity 25.9%; Pred. No. 8.2e-33;

Matches 170; Conservative 106; Mismatches 249; Indels 132; Gaps 18;

```

OY 30 ELGFHNNNIKAIFEKAFMGNPLQTHFYDNPT-QFVGRSAFOYLKRLHTLSNGAMDIO 88
Db 49 ELGFVYTKRAIDPKGAFSGDGLKEIEISONDVLEAVWFSLPRLHEIRIKANNLL 108
OY 89 EF-PDLKGTSLSEITLTRAGIRLLPSGMCQOLPRLVLELSHNQIELDLSLRQCKLEE 147
Db 109 YIDPD-----AFQNPRLRYLLISNGVAKHADVHRKIQSLQK 145
OY 148 T--GLQHN-RIMEIGADTFSSQLQALDLSWNAIRSIHPEARS-----TL 190
Db 146 VLLDIDQININITHVERNSFMGLSFESEMIWLMSKNGREINCAFNQGDLEMLSDNDL 205
OY 191 HSL-----VKLDLTDNQTLTLPLAGLGGIMHLKLGKALISQAQSKDPSFKRLTI 239

```

```

Db 206 EELPNDVPHGASGPVILLISRTIRHSLPSYGLLENKILKLNARSTYNLKLPTLEKEVTLME 265
Oy 240 LEVPAVYCCCPGCMGASFEFKASGOMEADLH-----LDDE 274
Db 266 ASLTPSHCCAF-----AMKROISDLHPICNKSILROEVDMVTOARGORVSLAEQG 317
Oy 275 ESSKRPGLGLARQENHYDQDLDELQLEMEDSKPHPSVOCSPPTGPKPCXYLFESWGIR 334
Db 318 ESS-----LAKEPTMSEFNVDLCNEVVD-----VICSPKPDAPNCECDIMGDILR 365
Oy 335 LAWVAIVLISVLCNGILVLTVPAGCAPLPVPKFEVGAAGANTLTGICGLASVDAIT 394
Db 366 VLIMEFSLIAIAGN-IIEIVLITTSQYKLTVPFLMCNLAFTDLIGIYLLIASIDHT 424
Oy 395 FQOFSEYGRMFEFGICRATGFLAVGSEASVILLTLAA-----VOCSSVY-- 440
Db 425 KSOYHNHYAINMOTGAGCDAAGFFTFASLSYTTLTATLERHHTTTAMQLOCKQVIRH 484
Oy 441 -SCVRAYGKSPSLGVSRAVGLCLALAGLAALPLASVGEYGAAPLCPYPAPPEGOPAL 499
Db 485 AASIMLYG-----WIFAFVAFVLPFGISSYMKVSICLPW--DIDSPLSQ 527
Oy 500 GTTVALVMNNSCFELVAVAGATTKLYCDLPRGD- FEAVWDCAMVRHVAWLIFADGLYCPY 558
Db 528 LTVVSLVLYLVNLAFAVYICGYTHIYLVNPNIMSSSDTKIAKRAMLIFTDPLCMVPI 587
Oy 559 AFLPSFAMGLPFPVPEAVKSVLLVPLPACINPLLYLFNFHPRDRLRLPRAG 615
Db 588 SFPAISASLKPVLITYSKILLVLFYPINSCANPFLYAFITKFRDVFILLSKFG 644

```

RESULT 4

145896

follicle stimulating hormone receptor - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 15-Oct-1996 #sequence_revision 15-Oct-1996 #text_change 21-Jan-2000

C:Accession: I45896

R:House, A.; Lambert, A.; Saumande, J.; Silversides, D.W.; Lussier, J.G.

Mol. Reprod. Dev. 39, 127-135, 1994

A:Title: Structure of the bovine follicle-stimulating hormone receptor complementary DNA

A:Reference number: I45896; MUID:951277199; PMID:7826612

A:Accession: I45896

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-695 <MUD>

A:Cross-references: GB:I22319; NID:g404671; PIDN:AAC37324.1; PID:g404672

C:Genetics:

A:Gene: FSHR

C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat F;71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 13.7%; Score 526.5; DB 2; Length 695;

Best Local Similarity 26.1%; Pred. No. 3.1e-32;

Matches 170; Conservative 108; Mismatches 290; Indels 83; Gaps 15;

```

Oy 30 ELGFHNNNIKAIPKAFMGNPLQTHFYDNP-I-QFVGRSAFOYLPRKLTSLNGANDIQ 88
Db 50 ELRFVLTKLVIRKAFSGGDELEKIEISQNDVLEAVFNSLPRKHEIRIKANML 109
Oy 89 EF-PDKGTTSLTLITTRAGIRLLPSGMCOQLPRRLVLESHNOIEELPSLHRCQLEE 147
Db 110 YIDPD-----AFONLPNRLYLLISNGIKHDAVHKIQSLQK 146
Oy 148 I--GLOHN-RIMEIGADTFQOLS-SLOALDSWNAIRSIHPEAF-----STL 190
Db 147 VLIDIDQINIHVERKSPFGLSPESMTWLSKNGIOIHNCANNGQDLDELNSDSNL 206
Oy 191 HSL-----VKLDITDNOITLTLPLAGLGLMLKLGKLNALSQAESKDSFKLRI 239
Db 207 EELPNDVFGASGVILDISRTIRSLPSYGLLENKILKLNARSTYNLKLPTLEKEVTLVE 266
Oy 240 LEVPAVYCCCPGCMGASFEFKASGOMEADLH-----LDDESSKRPLGLLARQA 288

```

```

Db 267 ASLTPSHCCAF-----AMKROISDLHPICNKSILROEVDMVTOARGORVSLAEED 318
Oy 289 ENHYDQDLDELQLEMEDSKPH--SVOCSPPTGPKPCXYLFESWGIRLAWVAIVLSVL 346
Db 319 EPSYAGFVPMSEFNVDLCNEVVDVTCSPEDPAFNPCEDIMGDILRLVIMEFSLIAT 378
Oy 347 CNGVLLTVFAGCAPLPVPKFEVGAAGANTLTGICGLASVDAITGQFSEYGRME 406
Db 379 GNILVLY-ILITSQYKLTVPFLMCNLAFTDLIGIYLLIASVDMVTKFEVHNVAIDMQ 437
Oy 407 TGIGCRATGFLAVGSEASVILLTLAAVOCSSVSGVRAYGKSPSLGVSRAVGLCLATA 466
Db 438 TGAGCDAAGFFTFASLSYTTLTATLERHHTTTAMQLOCKQVIRHVAWLIFADGLYCPY 497
Oy 467 GLAALPLASVGEYGAAPLCPYPAPPEGOPALGFTVALVMNNSCFELVAVAGATKYCD 526
Db 498 FAVALPFGISSYMKVSICLPW--DIDSPLSQYVMSLVLYLVNLAFAVYICGYTHIYLT 555
Oy 527 LPRGDF-EAVWDCAMVRHVAWLIFADGLYCPYAFLSFASMLLFPVTPRAVSVLLVYL 585
Db 556 VRPNITSSSDTKIAKRAMLIFTDPLCMAPISFPAISASLKVPLITYSKSKILLVLFY 615
Oy 586 PLPACINPLLYLFNFHPRDRLRLPRAGDSGPLAVAAAGELKSSCDST 636
Db 616 PINSCANPFLYAFITKFRDVFILLSKFG-----CYEQVQOTYRSTST 661

```

RESULT 5

JC1493

follicle stimulating hormone receptor - sheep

N:Alternate names: follicle stimulating hormone receptor

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jan-2000

C:Accession: JC1493; I47080

R:Khan, H.; Yarney, T.A.; Saitam, M.R.

Biochem. Biophys. Res. Commun. 190, 888-894, 1993

A:Title: Cloning of alternatively spliced mRNA transcripts coding for variants of ovine

A:Reference number: JC1493; MUID:93176195; PMID:8439338

A:Accession: JC1493

A:Molecule type: mRNA

A:Residues: 1-695 <KHA>

A:Experimental source: testis

R:Yarney, T.A.; Saitam, M.R.; Khan, H.; Rayndranath, N.; Payne, S.; Seidah, N.G.

Mol. Cell. Endocrinol. 93, 219-226, 1993

A:Title: Molecular cloning and expression of the ovine testicular follicle stimulating

A:Reference number: I47080; MUID:93351750; PMID:8394255

A:Accession: I47080

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-695 <YAR>

A:Cross-references: GB:L07302; NID:g165884; PIDN:AAA31525.1; PID:g165885

C:Genetics:

A:Gene: FSH-R

C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat F;71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 13.6%; Score 524.5; DB 2; Length 695;

Best Local Similarity 26.4%; Pred. No. 4.4e-32;

Matches 175; Conservative 101; Mismatches 273; Indels 113; Gaps 18;

```

Oy 30 ELGFHNNNIKAIPKAFMGNPLQTHFYDNP-I-QFVGRSAFOYLPRKLTSLNGANDIQ 88
Db 50 ELRFVLTKLVIRKAFSGGDELEKIEISQNDVLEAVFNSLPRKHEIRIKANML 109
Oy 89 EF-PDKGTTSLTLITTRAGIRLLPSGMCOQLPRRLVLESHNOIEELPSLHRCQLEE 147
Db 110 YIDPD-----AFONLPNRLYLLISNGIKHDAVHKIQSLQK 146
Oy 148 I--GLOHN-RIMEIGADTFQOLS-SLOALDSWNAIRSIHPEAF-----STL 190

```

Db 147 VLDDIQDNINITHVERNSFMGLSESMIVLMSKNGIOEIHNCANGTOLDLDELNSNL 206
 QY 191 HSL-----VKLDIDNQTLTTLPLAGLGMHLKLGKGNLALSOAFKSDSPFKRLI 239
 Db 207 EELPNDVFOGASGVILDISRTIRLSPLSGLENLKLKRASTYMLKLPISLEKEVTVE 266
 QY 240 LEVPYAVOCCPYGMCASFKAAGOMEADLH-----LDDESSKRLGLGLARQA 288
 Db 267 ASLTPSHCCAF-----AMNRQTSDDLHPICNKSILRQEVDMTQARGRISLAEDD 318
 QY 289 ENHYDDQDLDELQLEMDS--KPHPSVOCSPTPGPKCEYIPESSGIRLAWAIVLVL 346
 Db 319 EPSTYAGFDMYSEEDYDLCEVVVDVTCSEPDAFNCPEDIMGYDILVLIWFISILAIT 378
 QY 347 CNGLVLTIVFAGGAPAPLPVKFVVGAGTAGANTLTGISGLASVDALTFGQSESGARE 406
 Db 379 GNTLVIV-LITSOYKLTIVREFLMCNLAFAADLCIGIYILLASVDVHTKSOYHNVAIDM 437
 QY 407 TGLGCRATGFLAVLGSEASVLLTLAAVQ-----CSVSVSC--VRAVYKSPSLGSVRA 457
 Db 438 TGAGCDAAGFFTFVASELSVYTLTAITLERWHTITHAMQLECKYVHRHAASIMLVGW-- 495
 QY 458 GVLGCLALAGLAAALPLASVGEYASPLCLPYAPREGOPALGFTVALVMNSFCFLVVA 517
 Db 496 -----FAFVALFPLFGTISSTKVSICLPM--DIDSPISQLYVMSLLVNLVAFVVIC 546
 QY 518 GAYIKLYCDLPRGDF-EAVWDCAMVRHVAMLIFFADGLYCVAFLSFASMLGFPVTEBEA 576
 Db 547 GCYTHIYLVFRNPNITSSSDTKIAKRMAMLIITFDPLCMAPISFPAISASIKVPLITYSK 606
 QY 577 VASVLLVPLPACLPNLYLLENPHFRDLRLRPRAGDSGLAYAAAGLEKSSCDST 636
 Db 607 SKLLVLEFPINCANPFLAIFTKRRRDEFIL-----LSKRGCEYEV 649
 QY 637 QA 638
 Db 650 QA 651

RESULT 6

JC2237

Follicotropin receptor, testis - horse

N:Alternate names: eFSHR

C:Species: Equus caballus (domestic horse)

C>Date: 28-Aug-1985 #sequence, revision 07-Oct-1994 #text_change 13-Aug-1999

C:Accession: JC2237; JC2370

R:Robert, P.; Amsellem, S.; Christophe, S.; Benifla, J.L.; Bellet, D.; Koman, A.; Bizard

Biochem. Biophys. Res. Commun. 201, 201-207, 1994

A:Title: Cloning and sequencing of the equine testicular follicotropin receptor.

A:Reference number: JC2237; MUID:94256980; PMID:8198575

A:Accession: JC2237

A:Molecule type: mRNA

A:Residues: 1-694 <ROB>

A:Cross-references: GB:S70150; NID:9546896; PIDN:AA30854.1; PID:9546897

A:Experimental source: testis

C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat

C:Keywords: glycoprotein; hormone receptor; transmembrane protein

F:56-70/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F:71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F:96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F:121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>

F:146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

F:172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>

F:194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

F:366-386/Domain: transmembrane #status predicted <TM1>

F:398-440/Domain: transmembrane #status predicted <TM2>

F:443-464/Domain: transmembrane #status predicted <TM3>

F:485-507/Domain: transmembrane #status predicted <TM4>

F:528-549/Domain: transmembrane #status predicted <TM5>

F:573-596/Domain: transmembrane #status predicted <TM6>

F:608-629/Domain: transmembrane #status predicted <TM7>

F:191,199,268,293/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.6%; Score 522; DB 2; Length 694;
 Best Local Similarity 25.6%; Pred. No. 6,9e-32;
 Matches 168; Conservative 113; Mismatches 282; Indels 92; Gaps 15;

QY 30 ELGFHNNNIKAIEKAFMGWPLLOTIFYDNP1-QFGRSAFOYLPRLHTLSLNGADIQ 88
 Db 50 ELRFVTLKLRVIRKGAFFSGGDLKEIISQNDVLEIVANFSLPRLHRIEIKANN-- 107
 QY 89 EEPDLKGTTSLETLTLTRAGIRLLPSGMCQQLPRLVRLIELSHNOIEELPSLRHCQLEEI 148
 Db 108 -----LXYIDHDAFQNLPLMQLYLLISNTGIRKLPAVHKIQSLQKV 147
 QY 149 --GLOHN-RIMEIGADTFEQLS-SIQALDSWNAIRSIHEAF-----STLH 191
 Db 148 LDDIQDNINITHVERNSFMGLSESTTLRISKNGIOEIHNCANGTOLDLDELNSYNL 207
 QY 192 ST-----VKLDIDNQTLTTLPLAGLGMHLKLGKGNLALSOAFKSDSPFKRLI 240
 Db 208 EELPNDVFOGASGVILDISGTRHSLPNVGLLENLKLKRASTYMLKLPISLEKEVALMEA 267
 QY 241 EVPYAVOCCPYGMCASFKAAG-----QMEADLHLDDESSKRLPG 282
 Db 268 NLTPSHCCAF--AMNRQTSSELQTCNKSILROEYDMTQARGRVSIAEDDESSYKPG 324
 QY 283 LLARQENHYDDQDLDELQLEMDSKPHPSVOCSPTPGPKCEYIPESSGIRLAWAIVL 342
 Db 325 FDMMYSEFYD-----LCNEVD-----VTCSPKPDAPNCPEDIMGYDILRVLIWFIST 373
 QY 343 LSVLCLGLVLTIVFAGGAPAPLPVKFVVGAGIAGANTLTGISGLASVDALTFGQSESEY 402
 Db 374 LAITGN-IIVLVILITSOYKLTIVREFLMCNLAFAADLCIGIYILLASVDVHTKSOYHNVA 432
 QY 403 ARWETGLGCRATGFLAVLGSEASVLLTLAAVQCSVSVSVRAVYKSPSLGSVAVLGC 462
 Db 433 IDWGTGAGCDAAGFFTFVASELSVYTLTAITLERWHTITHAMQLECKYVHRHAASVWLVG 492
 QY 463 LALAGLAAALPLASVGEYASPLCLPYAPREGOPALGFTVALVMNSFCFLVAVAGYIK 522
 Db 493 WIFAFVALLPFEGISTYKVSICLPM--DIDSPISQLYVMSLLVNLVAFVVICGYIH 550
 QY 523 LYCDLPRGDF-EAVWDCAMVRHVAMLIFFADGLYCVAFLSFASMLGFPVTEBAVKSVL 581
 Db 551 YLTVRNPVIVSSSDTKIAKRMALITFDPLCMAPISFPAISASIKVPLITYSKILL 610
 QY 582 LVPLPACLPNLYLLENPHFRDLRLRPRAGDSGLAYAAAGLEKSSCDST 636
 Db 611 VLFTYINSCANPFLAIFTKRRRDEFILSKFG-----CYEMQAOQLVTFETIST 660

RESULT 7

A34548

Follictropin receptor precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 22-Jan-1993 #sequence, revision 22-Jan-1993 #text_change 13-Aug-1999

C:Accession: A34548; A41729

R:Sprengel, R.; Braun, T.; Nikolics, K.; Segaloff, D.L.; Seeburg, P.H.

Mol. Endocrinol. 4, 525-530, 1990

A:Title: The testicular receptor for follicle stimulating hormone: structure and func

A:Reference number: A34548; MUID:91125358; PMID:2126341

A:Accession: A34548

A:Molecule type: mRNA

A:Residues: 1-692 <SPR>

A:Cross-references: GB:102842; NID:g204183; PIDN:AAA41175.1; PID:g204184

R:Heckert, L.L.; Daley, L.J.; Griswold, M.D.

Mol. Endocrinol. 6, 70-80, 1992

A:Title: Structural organization of the follicle-stimulating hormone receptor gene.

A:Accession number: A41729; MUID:92149579; PMID:1738373

A:Reference: A41729

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-692 <HMC>

A:Cross-references: GB:581198; NID:g245344; PIDN:AA321415.1; PID:g245345

A>Note: sequence inconsistent with the nucleotide translation

QY 241 EPPYAYCCPCYMGCAFSPKASGOWEADLH-----LDDEESKREPLGLLARQAE 289
 Db 268 SLTPSPSCAF-----AMRROISELHPICRKSJLROEVDMTQRCGRSLAEDE 319
 QY 290 NHTDQDIDELQLEDESKRPH--SVQSPPTGPKPCEYLFESMGIRLAWAYLLSLVC 347
 Db 320 SSYSRQDMTYAEFDYDLCEWVDVYTCSPKDAFNPCEDILGNYLLVLLFISLITG 379
 QY 348 NGVLVLFVFGFAPLPVPEVVGATGANTLGLISGGLASDALTFGGFSEYGARME 407
 Db 380 NIIVVTL-ITTSQYKLVLPFELMCNLAFAADLCIGIYLLLASVDIHTKSOYHNVAIDMQ 438
 QY 408 GLGCRATGLFVAVGSEASVLLTLAAVQ-----CSVSCVAVRAGKSPSLGSV 455
 Db 439 GACCCDAGFEVYVASELSTVTLATLERNHTITHAMQDCKHVR-----HAA 487
 QY 456 RAGVLGCLAGLAALPLASVGEYASPLCELPYAPPEGOPALGFTVALVMNSCEFLV 515
 Db 488 SVMWGMW-FAPRAALPPLFGISSYMKVSIQLEM--DIDPSLQLYVMSLIVNLAFV 544
 QY 516 VAGAVIKLYCDLPGRDF-EAVWDCAMVRHVAWLIFADGLLYCVAVLSPASMLGLEPVP 574
 Db 545 ICGCYTHLYLVFNRPVIVSSSDTRAKRMALITDFLCMAPISEFAISASKVPLTV 604
 QY 575 EAVKSVLLVLPPLACINLTLLENPHFRDRLRLPRAGDSGLPVAAGELKESCD 634
 Db 605 SKAKILLVLPYINSCANDELYAIFTKNRRDFILLSKFG-----CYEMQOIVRTETS 659
 QY 635 ST 636
 Db 660 ST 661

RESULT 9

ORFUT

follicle receptor precursor - human
 N:Alternate names: follicle stimulating hormone receptor (FSHR)
 C:Contents: follicle receptor precursor long splice form; follicle receptor precursor
 C:Species: Homo sapiens (man)
 C:Date: 30-Sep-1991 #sequence-revision 06-Sep-1996 #text-change 22-Jun-1999
 R:Gromoll, J., Dankbar, B., Gudermann, T.
 Moll. Cell. Endocrinol. 102, 93-102, 1994
 A:Title: Characterization of the 5' flanking region of the human follicle-stimulating hormone receptor
 A:Reference number: 157661; MUID:95011044; PMID:7926278
 A:Accession: 157661
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-51 <GRK>
 R:Gromoll, J., Ried, T., Holtgreve-Grez, H., Nieschlag, E.; Gudermann, T.
 J. Mol. Endocrinol. 12, 265-271, 1994
 A:Title: Localization of the human FSH receptor to chromosome 2 p21 using a genomic probe
 A:Reference number: 155448; MUID:95000244; PMID:7916967
 A:Accession: 155448
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 286-695 <GR2>
 R:Gromoll, J., Gudermann, T., Nieschlag, E.
 Biochem. Biophys. Res. Commun. 188, 1077-1083, 1992
 A:Title: Molecular cloning of a truncated isoform of the human follicle stimulating hormone receptor
 A:Accession: PC1147
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-223, 286-294, 'P', 296-342 <GR3>
 A:Cross-references: EMBL:X68044; NID:931473; PIDN:CAA48179.1; PID:931474
 A:Experimental source: testis
 R:Gromoll, J.
 submitted to the EMBL Data Library, August 1992
 A:Reference number: S30560

A:Accession: S30560
 A:Molecule type: mRNA
 A:Residues: 1-12, 'R', 14-223, 286-294, 'P', 296-342 <GR4>
 A:Cross-references: EMBL:X68044; NID:931473; PIDN:CAA48179.1; PID:931474
 R:Kelton, C.A.; Cheng, S.V.; Nugent, N.P.; Schweichardt, R.L.; Rosenthal, J.L.; Over
 Mol. Cell. Endocrinol. 89, 141-151, 1992
 A:Title: The cloning of the human follicle stimulating hormone receptor and its expression
 A:Reference number: 157672; MUID:93246012; PMID:1301382
 A:Accession: 157672
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-679, 'N', 681-695 <KEL>
 R:Minagishi, T.; Nakamura, K.; Takakura, Y.; Ito, Y.; Igarashi, M.
 Biochem. Biophys. Res. Commun. 175, 1125-1130, 1991
 A:Title: Cloning and sequencing of human FSH receptor cDNA.
 A:Reference number: JN0122; MUID:91222171; PMID:1709010
 A:Accession: JN0122
 A:Molecule type: mRNA
 A:Residues: 1-111, 'T', 113-196, 'AV', 199-306, 'A', 308-695 <MIN>
 A:Cross-references: EMBL:M65085; NID:9182770; PIDN:AAA52477.1; PID:9182771
 C:Genetics:
 A:Gene: GDB:FSHR
 A:Map position: 2p21-2p16
 A:Introns: 223/3
 A:Note: The exact position of the intron cannot be determined from the experimental data
 C:Function:
 A:Description: receptor that mediates the biochemical effects of follicle stimulating hormone
 C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat
 C:Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone receptor
 F:1-695/Product: follicle receptor precursor, long splice form #status predicted <GR1>
 F:1-223, 286-695/Product: follicle receptor precursor, short splice form #status predicted <GR2>
 F:1-15/Domain: signal sequence #status predicted <SIG>
 F:16-695/Product: follicle receptor #status predicted <GR3>
 F:16-366/Domain: extracellular hormone binding #status predicted <EHB>
 F:71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
 F:96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
 F:121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F:146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
 F:172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
 F:194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
 F:367-387/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
 F:398-421/Domain: transmembrane #status predicted <TM1>
 F:444-465/Domain: transmembrane #status predicted <TM2>
 F:486-508/Domain: transmembrane #status predicted <TM3>
 F:529-550/Domain: transmembrane #status predicted <TM4>
 F:574-597/Domain: transmembrane #status predicted <TM5>
 F:609-630/Domain: transmembrane #status predicted <TM6>
 F:191, 199, 293, 318/Binding site: carbohydrate #status predicted <CAR>
 F:555/Binding site: phosphate (Thr) (covalent) #status predicted
 F:556/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
 Query Match 12.9%; Score 498.5; DB 1; Length 695;
 Best Local Similarity 24.8%; Pred. No. 4, 4e-30;
 Matches 163; Conservative 116; Mismatches 284; Indels 93; Gaps 18;
 QY 30 ELGFHNNNKAIPKEMGNPLQTHFYDNP-QLYGRSAPFLPKLHTLSUNGADIQ 88
 Db 50 ELRFVLTKLAVIVKAGFSGGDELEKIEISONDVLEADVFNQLPRLEIRIEKANNL 108
 QY 89 EPPDLKGTSLTLTLFRAGIRLLPSGMCQPLRLVRLVLSHNOIEPLSLHRCQKLEI 148
 Db 109 -----LYINPEAF-QNLPNLYGLISNGIKHLPVHRHISLQKV 147
 QY 149 --GLQHN-RIMEGADTFEQLS-SIQALDISWNAIRSHPEAF-----STH 191
 Db 148 LLDIDQDNIHITERSFVLSFESVILWLNKNGIQRHNCALFNGQDELINSDNNNE 207
 QY 192 ST-----VKLDLDNQTLTPLAGLGMHLKLGKLAISQAFSKSPKRLTL 240
 Db 208 ELPNDFHAGASPVILDISRTIRHSPYSGLENLKKLRARSTYNNKLPLEKVALMBA 267

```

QY 241 EVPRAYOCCEPYGMCASFEPKASGOMEADLH-----LDDESSKRPLGLLARQAE 289
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 268 SLATPSHCCAF-----ANMRQISFLHPICKSLIRQGVDMYQTRGQRSSLAEDNE 319
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 290 NHHYDODDELQLEMEDSKRPH--SVQCSPTPGPKPCCEYLFEWSGIRLAWAIVLSTLC 347
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 320 SSSYRGFMTEPTEPDYDLCEVVDVTCSPKPDAPNCPEDIMAGYNLRLVLIWFISTLATG 379
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 348 NGVLVLFVYAGGPAPLPVYKVVGAIAAGANTLGTISGLLSVDALTFGQSEYGARMET 407
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 380 N-IIVLVLTTSQYKLTVPREFLCMLAFADLCIGIYLLIASVDLHTHSQYINVAIDMQT 438
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 408 GLGCATGFLVILGSEASVLLTLTAAYQ-----CSVSVSCVYAVGKSPSGSVRAGVLG 461
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 439 GAGCGAAEFFYFVASELSYTLTALTLEKMHITTHAMQIDC-----KVQLRHASVYWG 493
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 462 CLIALAGLAALPLASVGEYASPLCLPYAPPGOPALGFTVALVMMNSFCELVVAGAYI 521
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 494 WI-FAFAALPFIPIGSISSYMKVSIQIPM--DIDSPLSQLYVMSLVLNLAIVVIGCXYI 550
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 522 KLYCDLPKQDF-EAVWDCAMVRHVAMLIFADGLLYCPVAFSPASMLGLPYVPEAVKSV 580
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 551 HLYLVARNPNIVSSSDPTIAKRMAMLIFTDFLCMAPISFFAISASLAKVPLITVSKATIL 610
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 581 LVLVLPPLACLPNLLYLLFNPHFRDDLRLRRPAGSGPLVAAAGLEKSSQDST 636
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 611 LVLFHPINSCANPFLVAFITKFRDFILLSKG-----CYEMQOIYRTETST 661
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
folliotropin receptor precursor - newt
N:Alternate names: follicle-stimulating hormone receptor
C:Species: Cynops pyrrhogaster (newt)
C>Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 31-Dec-2000
C:Accession: J07361
R:Nakayama, Y.; Yamamoto, T.; Oba, Y.; Nagahama, Y.; Abe, S.
Biochem. Biophys. Res. Commun. 275, 121-128, 2000
A:Title: Molecular cloning, functional characterization, and gene expression of a foll
A:Reference number: J07361
A:Contents: Testis
A:Accession: J07361
A:Molecule type: mRNA
A:Residues: 1-696 <NA>
A:Cross-references: DBJ:AB005587
C:Comment: This protein, containing seven transmembrane domains and a large glycosylated
in and thyroid stimulating hormone. This receptor has a common signal transduction pathw
C:Genetics:
A:Gene: fsh-r
A:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
C:Keywords: disulfide bond; glycolysis; glycoprotein; hormone receptor; testis; transmem
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-696/Product: follicle-stimulating hormone receptor #status predicted <MAT>
F:18-359/Domain: extracellular #status predicted <TM1>
F:370-389/Domain: transmembrane #status predicted <TM2>
F:402-424/Domain: transmembrane #status predicted <TM3>
F:447-468/Domain: transmembrane #status predicted <TM4>
F:488-511/Domain: transmembrane #status predicted <TM5>
F:532-553/Domain: transmembrane #status predicted <TM6>
F:577-600/Domain: transmembrane #status predicted <TM7>
F:612-653/Domain: transmembrane #status predicted <TM8>
F:616-190,198,267,292/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:445-520/Disulfide bonds: #status predicted

Query Match 12.0%; Score 463; DB 2; Length 696;
Best Local Similarity 24.8%; Pred. No. 2,3e-27;
Matches 158; Conservative 109; Mismatches 283; Indels 86; Gaps 17;

```

```

QY 89 EF-PD-LKGTSLTLTLTRAGIRLPLSGMCOQDRLRLVLELSHNOIELDPSLR--CQK 144
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 109 YIDPDAFONLPISKLLILSNIGQIVPA-----VSKIRSFHSLVDV 150
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 145 LEEIGLOHNRIMEIGADTPSOLSLQ-ALDLSMNAIRSHPEAFSTLSVLDLTQNO- 202
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 151 QDNINIRH-----ICKNSFAGLSSESITIRLNKNQIEEONHAFNGTH-LNLENLSDNR 204
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 203 -----LTTLPAGLIGIMHLKLNALSGAFSKDSEPKLR 238
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 205 LEKLDPOFGATGVYIIDISRTIRHFLPNNGIENIKKFRARFNYYLKKLPLEKFAELI 264
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 239 ILEVYAYOCCEYG-----MC-ASFKASGOMEADLH-----DDESSKR 280
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 265 EANLTVPSHCCAFANERKSEMHPICNKSPFGKHSAPKPEDEKNLRRESNEDLYSGRS 324
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 281 LGLARQAEHHYDODDELQLEMEDSKRPHSVQCSPTPGPKPCCEYLFEWSGIRLAWAIV 340
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 325 YSLVNGDEFNDYILCN--EVHD-----VICPKPDAPNCPEDIMGDNLTRLVLIWL 375
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 341 VLLSVLCNGVLTVEAGGPAPLPVYKVVGAIAAGANTLGTISGILLASVDALTFGQFSE 400
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 376 SLIATIGN-ITVLVILISQYKFTVPREFLCMLAFADLCIGIYLLIASVDITKSOYIN 434
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 401 YGARWETGLCCRTGFLAVLGSSEASVLLTLTAAYQCSVSVSCVYAVGKSPSGSVRAGVL 460
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 435 HAIDWQFGSGCAAGFTVYASLSYTLTITVITERMHTTYAMQDRKVRFRHATIMV 494
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 461 GCLIALAGLAALPLASVGEYASPLCLPYAPPGOPALGFTVALVMMNSFCELVVAGAY 520
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 495 VGMIFAFVAILPFIQVSVTKVSIQIPM--DIESRLSOSIYFVILVNCAPFLIICACY 552
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
A40077
thyrotropin receptor precursor - dog
N:Alternate names: thyroid-stimulating hormone receptor; TSH receptor
C:Species: Canis lupus familiaris (dog)
C>Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 13-Aug-1999
C:Accession: A40077; S06933
R:Parmentier, M.; Libert, F.; Maenhaut, C.; Lefort, A.; Gerard, C.; Perret, J.; Van S
Science 246, 1620-1622, 1989
A:Title: Molecular cloning of the thyrotropin receptor.
A:Reference number: A40077; MUID:90084524; PMID:2556796
A:Accession: A40077
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-764 <PA>
A:Cross-references: GB:M29957; NID:9164098; PIDN:AA30901.1; PID:9164099
R:Parmentier, M.; Libert, F.; Maenhaut, C.; Lefort, A.; Gerard, C.; Perret, J.; van S
Nucleic Acids Res. 17, 10493, 1989
A:Title: Nucleotide sequence of the dog thyrotropin receptor cDNA.
A:Reference number: S06933; MUID:90098886; PMID:2602159
A:Accession: S06933
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-764 <PA>
A:Cross-references: EMBL:X17146; NID:9849; PIDN:CA35026.1; PID:9850
C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repea
C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F:1-20/Domain: signal sequence #status predicted <MAT>
F:21-764/Product: thyrotropin receptor #status predicted <MAT>
F:53-76/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F:77-101/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F:102-126/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

```


A:Experimental source: hypothalamus
 A:Note: the sequences of residues 4-7 and 8-17, 4136-439 and 440-449 are interchanged in C:Comment: This receptor belongs to the guanine nucleotide binding protein (G-protein) c
 C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
 F:1-21/Domain: signal sequence #status predicted <S>
 F:23-76/Product: thyroid stimulating hormone receptor #status predicted <MAT>
 F:179-200/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>
 F:303-382/Domain: extracellular #status predicted <EXC>
 F:416-439/Domain: transmembrane #status predicted <TM1>
 F:451-474/Domain: transmembrane #status predicted <TM2>
 F:495-518/Domain: transmembrane #status predicted <TM3>
 F:538-561/Domain: transmembrane #status predicted <TM4>
 F:581-604/Domain: transmembrane #status predicted <TM5>
 F:627-650/Domain: transmembrane #status predicted <TM6>
 F:661-684/Domain: transmembrane #status predicted <TM7>

Query Match 11.9%; Score 456.5; DB 2; Length 764;
 Best Local Similarity 23.9%; Pred. No. 8.4e-27;
 Matches 167; Conservative 116; Mismatches 293; Indels 123; Gaps 20;

37 NIKALPEKAFKGNPLLOTTHFYDNPIDQFGRSAFQYLPKLTHTLSINGAMDIQ-----PP 91
 43 DIQRPLSP---PSTQTLKLETHHKTTPSRATFSLPMSIKYILSIDATLOQLSHSY 98
 92 DLKGTSTLEILTRAGIRLPSGMOQQLPRLVLELSHNOIEELPSLRCKLE----- 146
 99 NLSKVTHEIRN-TES-LTYIDSGALKEPLKFLGIFNTGTRVFDLTKRIVSTDVEFTIL 156
 147 EIGLOHNRWEIGADTFQSSLS-----LOALDLSMNA-IRS 181
 157 EL-TDNPYMTSPVANAFOGLSNETTLTKLYNNGETSIQGHAFNGTKLDVAVLKKNYLIV 215
 182 IPEAFSTLHS-LVKLDLTDNQLTTLPLAGLGIMHLKGNLALSQAFSKDFPKRLIL 240
 216 IQDPAAGVYSGEPLTDIDISTVTALPESKLEHLKELIANWTWTKLPJLSFELHTRA 275
 241 EPPVAYOCPPY-----GMCASF-KASQW----- 264
 276 DLSYPHSCAFKQKNIRGLISLMCNESIMGLRQKRSASALNPGFYOEEDLDGSA 335
 265 -----EADLH-----LDDESS-----KRLGLARQAEHNDODLDELQLE 302
 336 GYKENSKEFQDTHNSHYVFFEDQDEITFGQELKNPQETIQAPFNHDIYVCGSSE 395
 -303 MEDSKPSPVOCSPPTGPFKCEYLFESMGIRLAWAIVLVSCLNGVLVLTFAAGPAP 362
 396 M-----VCTPKSDEFNCPEDIMGYKRLRVVWVSLALLGNVFLVILLTSHKL 446
 363 LPPKVVGAIAAGANTLTIGISGLASVDALTFGQSEYGARWETGLGCRATGFLAVGS 422
 447 TVP-REFLMCNIAFADFCMGVLYLLIASVDLYTOSEYVNHADIMQTGPGCTAGFTVFAS 505
 423 EASVLLTLAAVOCSSVSCVRAVYKSPSLGSVRAGVGLGALAGLAALPLASVGEYA 482
 506 ELSVYTLVTLLEKWAITFAHMLDRKIRLMAHAVIMLGWCCFLALLPLVGISSYAK 565
 483 SPLCPAPPEGCPALGFTVALVMNSFCFLVAVGATIKLYCDLPR-----GDPEAVMD 537
 566 VSTICLPR-DTERPLALAIYIIVLLNIIAFIYACACVYKITYIVRPHYKPGD-----KD 619
 538 CAMVRHVAAILFADGLLYCPAFLSFASMLGLFPVTPPAVSKVLLVLPPLACNPLLYL 597
 620 TRIAKRAVAIVLTFDFCMAPISFYALSALMKNKPLITVNSKILLVLEPLNSCANPLYA 679
 598 LFNPHFRDDLRLRPRAGDSGPLAYAAAGLEKSSCDST 636
 680 IFTKAFQDVFMLLSKFGICKRQAQATRGQ-RVSSKNSST 717

RESULT 14
 I48882
 thyrotropin receptor precursor - mouse
 N:Alternate names: thyroid-stimulating hormone receptor; TSH receptor

C:Species: Mus musculus (house mouse)
 C:Date: 15-Mar-1996 #sequence revision 15-Mar-1996 #text-change 13-Aug-1999
 C:Accession: I48882
 C:Stain: S.A.; Oates, E.L.; Hall, C.R.; Grumbles, R.M.; Fernandez, L.M.; Taylor, N.A.
 Mol. Endocrinol. 8, 129-138, 1994
 A>Title: Identification of a point mutation in the thyrotropin receptor of the hyl/hy
 A:Reference number: A54271; MUIID:94224232; PMID:8170469
 A:Accession: I48882
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-764 <RES>
 A:Cross-references: EMBL:U02602; NID:9575923; PIDN:AAB60455.1; PID:9575924
 C:Genetics:
 A:Gene: TSHR
 C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repea
 C:Keywords: G protein-coupled receptor; transmembrane protein
 F:53-76/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
 F:77-101/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
 F:102-126/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F:127-151/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
 F:152-176/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
 F:179-200/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
 F:201-226/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

Query Match 11.8%; Score 454.5; DB 2; Length 764;
 Best Local Similarity 22.6%; Pred. No. 1.2e-26;
 Matches 173; Conservative 124; Mismatches 324; Indels 143; Gaps 20;

17 EPPVAYRTIGRL-----QELGFHNNIKAIPEKAFMGNPLQTHF-YDNPIDQFGRS 68
 36 DFRVTCCKELHRIPSLPSTQTLKLETHHKTTPSRATFSLPMSIKYILSIDATLOQLSEPH 95
 69 AFQYLPKLTHTLSINGAMDIQF-PD-LKGTSTLEILTRAGIRLPSGMOQQLPRLR-- 124
 96 SFYLSKMTHEIRNRSYLYIDPDALTELPLKFLGIFNTGTRVFDLTKRIVSTDVEFTIL 150
 125 -----VLELSHN-----QIEELPSLRCKLEIEIGLOHNRWEIGADTFQSSLSQALDLSW 176
 151 DIFPLEITDNPYMTSPVANAFOGLSNETTLTKLYNNGETSIQGHAFNGTKLDVAVLKKNYLIV 209
 177 NA-IRSIHPEAFSTLHS-LVKLDLTDNQLTTLPLAGLGIMHLKGNLALSQAFSKDSF 234
 210 NKLYLTAIDNDAFGVYSGPFLTDVSSVTALPESKLEHLKELIAKDTWTKLPJLSLSF 269
 235 PKLILEPPVAYOCPPY-----GMCASF----- 257
 270 LHLTRADLSYPSHCCAFKQKNIRGLISLMCNESIRNLQRKRSVNLKSPYIOVEED 329
 258 -----FKASQW-----EADLHLDDESSKRLGLARQAEHNDODL 296
 330 PGDSVGYKQNSKQOESPNSHYVFFEDQDEYVGGQELKNPQETIQAPFNHDIYVCGSSE 389
 297 DELQLEMEDSKPSPVOCSPPTGPFKCEYLFESMGIRLAWAIVLVSCLNGVLVLTVE 356
 390 -----CGDNEDMVCPTKPSDEFNCPEDIMGYKRLRVVWVSLALLGNIVFLITLL 440
 357 AGGPAPLPPKVVGAIAAGANTLTIGISGLASVDALTFGQSEYGARWETGLGCRATGF 416
 441 TSHKLTVP-REFLMCNIAFADFCMGVLYLLIASVDLYTHSEYVNHADIMQTGPGCTAGF 499
 417 LAVGEASVLLTLAAVOCSSVSCVRAVYKSPSLGSVRAGVGLGALAGLAALPLAS 476
 500 FTVASSELSTVTLVITLERWATFAHMLDRKIRLMAHAVIMLGWCCFLALLPLVWVG 559
 477 VGEYGASPLCPAPPEGCPALGFTVALVMNSFCFLVAVGATIKLYCDL-----PRG 530
 560 ISSYAKVSTICLPM-DTPTPLALAIYIIVLLNIIAFIYACACVYKITYIVRPHYKPGD 617
 531 DPEAVMDCAMVRHVAAILFADGLLYCPAFLSFASMLGLFPVTPPAVSKVLLVLPPLAC 590
 618 K-----DTKIAKRAVAIVLTFDFCMAPISFYALSALMKNKPLITVNSKILLVLEPLNSC 672
 591 LNPLLYLFNPHFRDDLRLRPRAGDSGPLAYAAAG-----ELEKSSCDSTQAL 639

Db 673 ANPFLAFTFAFQDQVLTLLSKFCICKRQAQAYQGRVCPNNSTIGIQLOKLPQDTRQSL 732
 QY 640 VAFSDVDLLELASEAGRPPELTGFPSTVLLISCOQPGAPRLG 683
 Db 733 PNMDDTYELLGNSQL-----APKIQG 753

RESULT 15

A42395

lutropin receptor - mouse

N/Alternate names: luteinizing hormone-choriogonadotropin receptor

C/Species: Mus musculus (house mouse)

C/Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 13-Aug-1999

C/Accession: A42395

R/Gudermann, T.; Birnbaumer, M.; Birnbaumer, L.

J. Biol. Chem. 267, 4479-4488, 1992

A/Title: Evidence for dual coupling of the murine luteinizing hormone receptor to adenyly

lone receptor expressed in L cells.

A/Reference number: A42395; MUID:92165799; PMID:1311310

A/Accession: A42395

A/Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-700 <GUD>

A/Cross references: GB:M81310; GB:M81318; NID:q198811; PIDN:AAA39432.1; PID:g198812

A/Note: Sequence extracted from NCBI backbone (NCBIN:84064, NCBI:84066)

C/Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h

C/Keywords: G protein-coupled receptor; transmembrane protein

F:54-77/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F:78-102/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F:103-127/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F:128-152/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>

F:153-177/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

F:180-201/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>

F:202-226/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

Query Match 11.7%; Score 449.5; DB 2; Length 700;

Best Local Similarity 23.18; Pred. No. 2.6e-26;

Matches 159; Conservative 124; Mismatches 306; Indels 99; Gaps 18;

QY 38 IKAIPKAFMG-NPLQTIHFYDNPQFVGSAPQYLPKLHTLSLNGAMDQEPFDLKG 96
 Db 65 VKVIFPQAFRGLNEVVKIEISQSDSLRIEANAFDNL-----LNL----- 105
 QY 97 TSELITLTRAGRLRPSGMCQOLPRLVLELHNOIEELPSLHR-----CQK 144
 Db 106 ---ETLIQNTKKNLTYEPGAFNMLPRKYLISICNTGIRTLPDVSKISSSEFFILEICDN 162
 QY 145 L-----EEIGIQ--HNRIMEIGADTFSQSLQALDLSMNA-IRSIPEA 186
 Db 163 LYITTPGNAFOGMNNESTITKLXNGFEVQSHARNG-TTLISLELKENITYLEKMHSGT 221
 QY 187 ESTLSLVKLDLTNOQLTLPAGLGLMHLKLGMLALSQAQSKDSFPKRLILEVPYAY 246
 Db 222 FQGATGPSILDVSTRLQALPSHGLESIQTLIATSSYSKTLPSREKFTSLVATLTYP 281
 QY 247 OCCPYGMCASFQAGQWAEIDLHDESSKRPGLIARQAEHN--YDQDLDELQ--- 301
 Db 282 HCCAFRNLPK-----KEQNFSEIFENFSKQCESTVREANNEETLYSAIFENELSGW 333
 QY 302 --EMEDSKRHPSPVQSPNPPKCEYLFEESWIRLAWAIVLSYLCNGLVLTVPAG 359
 Db 334 DYDYDFCSP-KTLQCTPERDAFNPCEIDINGVAFRLVLMINTALFNG-LTVLFLVLT 391
 QY 360 PAPLPVAFVVAIGANTLTISGLASVDALTFGQSEYGARWETGLGRATGFLAV 419
 Db 392 RYKLVPRFLKMCNLSFADFQCMGLYLLILASVDSQTKGYNHAIDWOTGSGCSAAGFTV 451
 QY 420 LGSEASVLLTFLAAYQCSVSVARAYGKSPISGSRAGVGLGCLALAGLAALPLASVGE 479
 Db 452 FASELSVYTLVITIERNHTITTYAVQLODKLRLRAIPIMLGWIPTLWATPLVGVSS 511
 QY 480 YGASPLCLPYAPPEGOPALGFTVALVMMNSFCFLVAVAGATIKLYCDLPRGDEAV-WDC 538

Db 512 YMKVSIQPM--DYESTLSQYIIISILLNNVAFVIGACVIRITFANQNEPLTPPNKDT 569
 QY 539 AMRVHVMILTFADGLLCPVAFLEFASMLGLEPVTPEAVKSVLLVLPAPCLNPLVLL 598
 Db 570 KIARKMALIFTDFTCMAPISFEFAISAFKVPILTVTNSKVLLVLFYVNSCANPEFLYAV 629
 QY 599 FNPFRDRLRLRPAGDSGPLAVAAAGELKESQDSTQALVAFSDVDLLELASEAGRP 658
 Db 630 FTKAFQDRDFLLLSRFG-----CCKHRAELIRRKRESACTPNSKNGFPR 673
 QY 659 GLETFGFP-----VTLISCOQPGAPRL 681
 Db 674 SSR----PSQAALKLSTVHCQOPPPRV 697

Search completed: November 8, 2002, 19:36:53
 Job time : 22.8167 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2002, 19:34:40 ; Search time 8.6019 Seconds
(without alignments)
1233.683 Million cell updates/sec

Title: US-09-851-595-8
Perfect score: 3850
Sequence: 1 GILHLETLIDLNKQLEFPV.....GGLSGGGGFPQSLAFASHV 736

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 92612 seqs, 14418503 residues 92612

Total number of hits satisfying chosen parameters:
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications, AA:*

- 1: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB pep:*
- 2: /cgn2_6/ptodata/1/pubppa/PCT_NEW_PUB pep:*
- 3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB pep:*
- 4: /cgn2_6/ptodata/1/pubppa/US06_PUB pep:*
- 5: /cgn2_6/ptodata/1/pubppa/US07_PUB pep:*
- 6: /cgn2_6/ptodata/1/pubppa/US07_PUB pep:*
- 7: /cgn2_6/ptodata/1/pubppa/US09_PUB pep:*
- 8: /cgn2_6/ptodata/1/pubppa/US09_PUB pep:*
- 9: /cgn2_6/ptodata/1/pubppa/US09_PUB pep:*
- 10: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB pep:*
- 11: /cgn2_6/ptodata/1/pubppa/US10_PUB pep:*
- 12: /cgn2_6/ptodata/1/pubppa/US10_PUB pep:*
- 13: /cgn2_6/ptodata/1/pubppa/US10_PUB pep:*
- 14: /cgn2_6/ptodata/1/pubppa/US10_PUB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1103	28.6	631	US-09-862-767A-2	Sequence 2, Appl
2	790	20.3	436	US-09-862-767A-4	Sequence 4, Appl
3	783	20.3	431	US-09-862-767A-8	Sequence 8, Appl
4	727.5	18.9	358	US-09-862-767A-6	Sequence 6, Appl
5	688.5	17.9	861	US-09-804-551B-20	Sequence 20, Appl
6	513.5	13.3	675	US-09-877-804-7	Sequence 7, Appl
7	513.5	13.3	695	US-09-877-804-6	Sequence 8, Appl
8	500.5	11.0	695	US-09-804-626-6	Sequence 3, Appl
9	443.5	11.5	674	US-09-877-804-3	Sequence 3, Appl
10	443.5	11.5	700	US-09-877-804-2	Sequence 2, Appl
11	425.5	11.1	718	US-09-928-175-21	Sequence 21, Appl
12	425.5	11.0	737	US-09-928-175-20	Sequence 20, Appl
13	423.5	11.0	737	US-09-804-626-6	Sequence 6, Appl
14	420	10.9	699	US-09-804-626-6	Sequence 8, Appl
15	410.5	10.7	730	US-09-928-175-7	Sequence 7, Appl
16	410.5	10.7	730	US-09-928-175-7	Sequence 13, Appl
17	396.5	10.3	646	US-09-928-175-13	Sequence 12, Appl
18	396.5	10.3	682	US-09-928-175-12	Sequence 12, Appl

20	389	10.1	757	10	US-09-928-175-24	Sequence 24, Appl
21	370.5	9.6	260	10	US-09-925-299-975	Sequence 975, Appl
22	370.5	9.6	458	10	US-09-862-767A-9	Sequence 9, Appl
23	350.5	9.1	359	10	US-09-862-767A-7	Sequence 7, Appl
24	319	8.3	458	10	US-09-862-767A-5	Sequence 5, Appl
25	226	5.9	605	10	US-09-782-980-62	Sequence 62, Appl
26	224.5	5.8	1059	9	US-09-905-291A-290	Sequence 290, Appl
27	224.5	5.8	1059	9	US-09-905-291A-290	Sequence 290, Appl
28	224.5	5.8	1119	9	US-09-905-291A-294	Sequence 294, Appl
29	224.5	5.8	1119	9	US-09-905-291A-294	Sequence 294, Appl
30	224.5	5.8	1119	10	US-09-905-291A-294	Sequence 294, Appl
31	224.5	5.8	1119	10	US-09-789-804-2	Sequence 2, Appl
32	220	5.7	794	10	US-09-877-804-8	Sequence 8, Appl
33	216.5	5.6	438	10	US-09-758-140-4	Sequence 4, Appl
34	216	5.6	473	10	US-09-796-858-36	Sequence 36, Appl
35	216	5.6	473	10	US-09-893-348-27	Sequence 27, Appl
36	216	5.6	473	10	US-09-972-599A-4	Sequence 4, Appl
37	212.5	5.5	640	9	US-09-905-291A-292	Sequence 292, Appl
38	212.5	5.5	640	9	US-09-905-291A-292	Sequence 292, Appl
39	212.5	5.5	640	10	US-09-989-722-501	Sequence 501, Appl
40	212.5	5.5	640	10	US-09-989-722-501	Sequence 501, Appl
41	212.5	5.5	640	10	US-09-989-722-501	Sequence 501, Appl
42	212.5	5.5	640	10	US-09-989-722-501	Sequence 501, Appl
43	212.5	5.5	640	10	US-09-989-731-501	Sequence 501, Appl
44	212.5	5.5	640	10	US-09-989-731-501	Sequence 501, Appl
45	212.5	5.5	640	10	US-09-989-732-501	Sequence 501, Appl

ALIGNMENTS

RESULT 1
US-09-862-767A-2
; Sequence 2, Application US/09862767A
; Patent No. US20020034786A1
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: TANGO-72 AND TANGO-77 NUCLEIC ACID MOLECULES AND POLYPEPTIDES
; FILE REFERENCE: MBIO1997-0081P(CIP)(M)
; CURRENT APPLICATION NUMBER: US/09/862,767A
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/127,856
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 60/054,646
; PRIOR FILING DATE: 1997-08-04
; SOFTWARE: FastSeq for Windows Version 3.0
; NUMBER OF SEQ ID NOS: 9
; SEQ ID NO 2
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-862-767A-2

Query Match 28.6%; Score 1103; DB 10; Length 631;
Best Local Similarity 40.1%; Pred No. 1.7e-78;
Matches 242; Conservative 96; Mismatches 209; Indels 56; Gaps 10;

QY	99	LELTITLFRAGIRLLPFGMCQQLPRRLVLELSHNOHELPSLRCKKLEIRIGLOHRIWEI	158
DB	1	LEHITLITGFISSIPNNLCQCKMLRTLLDSYNNINDLPSFGCHALEISIQNRIVYI	60
QY	159	GADTFSSLSQALDLSWNAIRSIHPEAFSTLSLVKLDLTQTLPLAGLGIMHLK	218
DB	61	KEGTFQGLISRLIDLSRNLHEHSRAFTLGPINLDVSPFELTSPFEGINGLNQK	120
QY	219	LKGNLALSOAFSKDPSFKRLRILEVPAVYVCCPYGMCAKFGKASGOWFAEDLHDDDESSK	278
DB	121	LVGNFKLELAADDFVNLKLSVPAVYVCCAFWGCDSY	173
QY	279	RPLGLAQAENHDDQDELQLEMEDSKPAP--VOCSPFGPFKPCERYLFESWGIRLA	356
DB	174	-----VADEKGTADANVTSTLENEL-HSQIITICTPSTGAFKCEVGLSGMWIRLT	224

```

Db      223 HREYASAPLCLEP--PTGETPSLGTVTLVLLNSIAFLIMAVITYIKVCONLEKEDLSSEN 280
QY      536 WDCAMRHYAVAMLIADGLLYCPVAFLSFASMLGLFPYTPENAKSVLLVYLPLPACINPLL 595
Db      281 SSSSMIKHYAMLIFTNCIFPCFPVAFSFAFLITTAISISEIMKSVTLFLFPACINPVL 340
QY      596 YLLFNPHEFDD---LRRLRPAGDSGLPATAAGBLEK-----SSCS 635
Db      341 IYEFNPKEKEDMKLLKRRYTKKSGSVSISISGGCLQDGFYDCGMVSHLQGNLTVDC 400
QY      636 TQALVAEFDV---LILPASEGRPPGLEIYGFPSVTLSCOOP 676
Db      401 CESFLTKRPVSCKHLI-----KSHSCPALAAVASCORP 432

RESULT 3
US-09-862-767A-8
; Sequence 8, Application US/09862767A
; Patent No. US20020034786A1
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: TANGO-72 AND TANGO-77 NUCLEIC ACID MOLECULES AND POLYPEPTIDES
; FILE REFERENCE: MBIO1997-008P1RCPLCN1(M)
; CURRENT APPLICATION NUMBER: US/09/862,767A
; CURRENT FILING DATE: 2001-05-21

```

```

; PRIOR APPLICATION NUMBER: US 60/054,646
; PRIOR FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
us-09-862-767A-8

Query Match
Best Local Similarity 20.3%; Score 783; DB 10; Length 431;
Matches 173; Conservative 73; Mismatches 157; Indels 60; Gaps 9;

QY 240 LEVYAYAOCCPYGACAFKASFGQMEADLHLD-----EESKRPGLILARQAEHNYDD 295
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2 LSVAYAYOCCAFWGCDSY-----ANLNTEDSLDHSVAOEGKGTADANVYSTLENEHSQ 57
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 296 LDELQLEMEDSKRPVSOCCPTPGPFKPCCYELFESWIGIRLAVNAIYLISVLCGVALTV 355
      : : : : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 58 I-----ILHCPSTGAEPFCEYICGWSWTRTPM----- : | | : : : : | | : | |

```

QY	356	FAGGAPAPLPVKKVYVGA	IAACANTLTGICGLLAS	VDALTFEGCFSEYGARETGLGRATG	411
Db	104	FACST-LEBSKPI	IGLITSYSLFMGIYTGIL	FLFDVAWSMGRAEIGIMWETGSGCKK	16
QY	416	FLAVLGSEASVLLTTLAA	OCSSVASCVRVYGGSP	SLGSVRAGVLCGLAAGLAAPLA	47
Db	163	FLAVSESSAIFELMLAT	VERBSLADIMKNGKSNHL	KOFRAALLAFLGAVACCFPAE	22
QY	476	SVGEYCASPLCIPYAP	BEQOPALGTTVALVMNS	FCLVLNAGAYIKYICDLPGRDDEAV	53
Db	223	HGEYSASPLCIP-	PGHEPTSLGFVTYLV	LNSLAFLMAIVYTKLYCNLEKDELSN	28
QY	536	WCCAAVRHVAAMLI	IFADGLLYCPVAELSF	ASMIIGLFPVPEAVKSVLLVLEPLPCINPL	59
Db	281	SOSSMKIHVAAMLI	FNCFCEFCVPAFSP	APLITAISSPELMKSVTLLEFPLPACLNPLV	34
QY	596	YLLFNPHERDD-	LRLTRRADDSGL	LAVAAAGELK-----SSCDS	635
Db	341	YVFENPKFEKML	KLKRVYTKKSSSV	SVSSISOGCEJODFYDCGMYSHLGMLTYCDC	400
QY	636	TQALVAFSDVD-	LLIFASEAGRPGL	ETVYGFPSVLLSCQO	675
Db	401	CESEFLTKRPV	SCKHLI-----KSSGDPALVAVSCOS	473	

Query Match	17.98;	Score 688.5;	DB 10;	Length 861;
-------------	--------	--------------	--------	-------------

RESULT 6
US-09-877-804-7
Sequence 7, Application US/09877804
Patent No. US20020061557A1
GENERAL INFORMATION:
APPLICANT: Nikolic, Karoly
APPLICANT: McFarland, Keith C.
APPLICANT: Seguro, Deborah L.
APPLICANT: Seeburg, Peter H.
TITLE OF INVENTION: Glycoprotein Hormone Receptor Molecules
FILE REFERENCE: P057691C2
CURRENT APPLICATION NUMBER: US/09/877,804
CURRENT FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: US 08/207,814
PRIOR FILING DATE: 1994-03-07
PRIOR APPLICATION NUMBER: US 07/781,153
PRIOR FILING DATE: 1991-10-31
PRIOR APPLICATION NUMBER: US 07/347,683
PRIOR FILING DATE: 1989-05-05
NUMBER OF SEQ ID NOS: 22
SEQ ID NO 7

LENGTH: 675
 TYPE: PR
 ORGANISM: Artificial sequence
 FEATURE:
 OTHER INFORMATION: deduced sequence
 US-09-877-804-7

Query Match 13.3%; Score 513.5; DB 10; Length 675;
 Best Local Similarity 25.9%; Pred. No. 2e-32;
 Matches 188; Conservative 106; Mismatches 284; Indels 147; Gaps 25;

```

QY 30 ELGFHNNNIKAPEKAFMGPNLLQTHFYDNP1-QFYGRSAFOYLPKLTLSLNGAMDIO 88
DB 33 ELRFVLTKLRIKRVIPKGSFAGGDLKEIISONDVLEIADVFSNLPKLEIRIEKANNL- 91
QY 89 EFPDCKGTTSLLELTLLTRAGIRLLPSGMCQOLPRLRYLELSHNOIELELPSLHRCQKLEI 148
DB 92 -----LYINPEAF-QNLPRLRYLLISNTGIKHLPVHKIQSLQKV 130
QY 149 --GLOHN-RIMEIGADTFESOLS-SLOALDSWNAISIHPEAF-----STLH 191
DB 131 LLDIDONINIHIAVARNFSGELSESVILMSKNGIEIHNCANFGTQDLDELNSDNNMLE 190
QY 192 SL-----VKLDLTNDQTLTLPPLAGLGLMHLKLGKGNLALSOAFKSDSEPKRLIL 240
DB 191 ELRPNDVFOGASGVILDISRTKTVHSLPNHGLENLKTLRARSTYRLKRLPNDKFTVLM EA 250
QY 241 EYFYAYOCCPYMGCAFSGFASGOWEADLH-----LDDEE 275
DB 251 SLTFPSHCCAF-----ANLKRQISELHPICNKSILRODIDDMTOIGDQVSLIDDEP 302
QY 276 SSKRPLGLLARQAEHNYDODLDELQEMEDSKPHPSVQCSPTPPGPKCEYLFEESMGIRL 335
DB 303 S-----YKGSMDMYNEFDYLCN---EYVD-----VTCSKPDAPNCEIDMGYNILRV 349
QY 336 AWAIVLLSVLCNGVLLTVFAGGAPRLPVKFEVYGALAGANTLTGISCGLASVDALTF 395
DB 350 LWFISILAITGNTTVLV-VLTTSQYKLTVPRLMCMNLAADLCIGIYLLIASVDIHTK 408
QY 396 GGFSEYGARWETGLGCRATGFLAVLGSEASVLLTLTAOQ-----CSVSVCVRAYGKS 449
DB 409 SOYHNVAIDMOTGACDAGGFFTVFASSELVYTLTATTLERWHTITIAMOLEC-----KV 463
QY 450 PSLGSVRAGVLCGLAALPLASVGEYASPLCLPYAPPEGAPALGFTVALVMNN 509
DB 464 QLRHNASVWVG-WTFEFAAALPFIIGISSYMKVSICLPM--DIDSPLSQLYVMALLVLN 520
QY 510 SFCFLVAVAGATIKLYCDLPRGDF-EAVWDCAMVRHVAWMLFEADGLLYCPVAFLSFASMLG 568
DB 521 VLAFFVYIGCYTHIYLTNRNPTIVSSSDTKIKRMAATLIFTDFLCMAPISFFAISASLK 580
QY 569 LFPVTPPEAVKSVLLVLLPLACLNPLLYLLENPHFRDRLRLRPRAAGSGPLAVAAAGEL 628
DB 581 VPLITVSKAKILLVLYFYNSCANPFLYALFTKNFRDFIL-----L 623
QY 629 EKSSCDSTQALVAFSDVDLLEASGRRPGLETYGFPSVTLLISCOQPGAPRLGESHGVE 688
DB 642 SFGCYEQMOAQIYRT-----ETSSA-----TINF-HARKSHCS--SAPRVNTSYVLV 667
QY 689 PEGNH 693
DB 668 PL-NH 671

```

RESULT 7
 US-09-877-804-6
 Sequence 6, Application US/09877804
 Patent No. US20020061557A1
 GENERAL INFORMATION:
 APPLICANT: Nikolic, Karoly
 APPLICANT: McFarland, Keith C.
 APPLICANT: Segal, Deborah L.
 APPLICANT: Seeburg, Peter H.

TITLE OF INVENTION: Glycoprotein Hormone Receptor Molecules
 FILE REFERENCE: P0576PIC2
 CURRENT APPLICATION NUMBER: US/09/877, 804
 CURRENT FILING DATE: 2001-09-04
 PRIOR APPLICATION NUMBER: US 08/207, 814
 PRIOR FILING DATE: 1994-03-07
 PRIOR APPLICATION NUMBER: US 07/781,153
 PRIOR FILING DATE: 1991-10-31
 PRIOR APPLICATION NUMBER: US 07/347,683
 PRIOR FILING DATE: 1989-05-05
 NUMBER OF SEQ ID NOS: 22
 SEQ ID NO 6
 LENGTH: 692
 TYPE: PR
 ORGANISM: Artificial sequence
 FEATURE:
 OTHER INFORMATION: deduced sequence
 US-09-877-804-6

Query Match 13.3%; Score 513.5; DB 10; Length 692;
 Best Local Similarity 25.9%; Pred. No. 2.1e-32;
 Matches 188; Conservative 106; Mismatches 284; Indels 147; Gaps 25;

```

QY 30 ELGFHNNNIKAPEKAFMGPNLLQTHFYDNP1-QFYGRSAFOYLPKLTLSLNGAMDIO 88
DB 50 ELRFVLTKLRIKRVIPKGSFAGGDLKEIISONDVLEIADVFSNLPKLEIRIEKANNL- 108
QY 89 EFPDCKGTTSLLELTLLTRAGIRLLPSGMCQOLPRLRYLELSHNOIELELPSLHRCQKLEI 148
DB 109 -----LYINPEAF-QNLPRLRYLLISNTGIKHLPVHKIQSLQKV 147
QY 149 --GLOHN-RIMEIGADTFESOLS-SLOALDSWNAISIHPEAF-----STLH 191
DB 148 LLDIDONINIHIAVARNFSGELSESVILMSKNGIEIHNCANFGTQDLDELNSDNNMLE 207
QY 192 SL-----VKLDLTNDQTLTLPPLAGLGLMHLKLGKGNLALSOAFKSDSEPKRLIL 240
DB 208 ELRPNDVFOGASGVILDISRTKTVHSLPNHGLENLKTLRARSTYRLKRLPNDKFTVLM EA 267
QY 241 EYFYAYOCCPYMGCAFSGFASGOWEADLH-----LDDEE 275
DB 268 SLTFPSHCCAF-----ANLKRQISELHPICNKSILRODIDDMTOIGDQVSLIDDEP 319
QY 276 SSKRPLGLLARQAEHNYDODLDELQEMEDSKPHPSVQCSPTPPGPKCEYLFEESMGIRL 335
DB 320 S-----YKGSMDMYNEFDYLCN---EYVD-----VTCSKPDAPNCEIDMGYNILRV 366
QY 336 AWAIVLLSVLCNGVLLTVFAGGAPRLPVKFEVYGALAGANTLTGISCGLASVDALTF 395
DB 367 LWFISILAITGNTTVLV-VLTTSQYKLTVPRLMCMNLAADLCIGIYLLIASVDIHTK 425
QY 396 GGFSEYGARWETGLGCRATGFLAVLGSEASVLLTLTAOQ-----CSVSVCVRAYGKS 449
DB 426 SOYHNVAIDMOTGACDAGGFFTVFASSELVYTLTATTLERWHTITIAMOLEC-----KV 480
QY 450 PSLGSVRAGVLCGLAALPLASVGEYASPLCLPYAPPEGAPALGFTVALVMNN 509
DB 481 QLRHNASVWVG-WTFEFAAALPFIIGISSYMKVSICLPM--DIDSPLSQLYVMALLVLN 537
QY 510 SFCFLVAVAGATIKLYCDLPRGDF-EAVWDCAMVRHVAWMLFEADGLLYCPVAFLSFASMLG 568
DB 538 VLAFFVYIGCYTHIYLTNRNPTIVSSSDTKIKRMAATLIFTDFLCMAPISFFAISASLK 597
QY 569 LFPVTPPEAVKSVLLVLLPLACLNPLLYLLENPHFRDRLRLRPRAAGSGPLAVAAAGEL 628
DB 598 VPLITVSKAKILLVLYFYNSCANPFLYALFTKNFRDFIL-----L 640
QY 629 EKSSCDSTQALVAFSDVDLLEASGRRPGLETYGFPSVTLLISCOQPGAPRLGESHGVE 688
DB 641 SKFGCYEQMOAQIYRT-----ETSSA-----TINF-HARKSHCS--SAPRVNTSYVLV 684
QY 689 PEGNH 693

```

RESULT 9

[illegible]

Db 425 DCLMGVYLFEVGIEDIKRYGQYKAYALLMESVQCRIMGLFAMLSTEVSVLLITLYLLEK 484
Qy 437 SVSV-----SCVRAKSPSLGSRVAGVGLCCLAG-LAAALPLASV-----GEYGASPLCL 487
Db 485 FLVIVFPESNIRPKGRQTSV-----ILICIMAGFLIAVIFPMNNDYGFNGKNGVCF 538
Qy 488 P-YAPPEGOPALGFTVALVM-MNSFCFLVAVAGYIKLYCDLPRGDPEA--VWDC----- 538
Db 539 PLYDQTEDIGSKGYSGLGFLGVNLAFLIIVFSYITMFCISQKTAQOTTEVANCGRREV 598
Qy 539 AMVRHVAWLIFADGLLYCPVAELFSPASMLGLFPV--TPRAVKS-VLLIVLPLPACINPLLY 596
Db 599 AVANRFFEFVSDAICWIPFVV---KILSLRVEIPDITMTSWIVIFPLVNSALNPILY 655
Qy 597 LLENPFRDDLRL 610
Db 656 TLTTFNFFDKLKL 669

RESULT 12
US-09-928-175-2
; Sequence 2, Application US/09928175
; Patent No. US20020123618A1

; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher J.
; APPLICANT: Gong, Jianhua
; APPLICANT: Daugherty, Betsy
; APPLICANT: Rogers, No. US20020123618A1ma
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 00-1229
; CURRENT APPLICATION NUMBER: US/09/928,175
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/224,455
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 754
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-928-175-2

Query Match 11.1%; Score 425.5; DB 10; Length 754;
Best Local Similarity 24.2%; Pred. No. 1.8e-25;
Matches 163; Conservative 105; Mismatches 235; Indels 171; Gaps 24;

Qy 3 HNEETLDLWYKLOEPP--VAIRTLGRLOELGFHNNNIKAIPKAFMGNPLLOTIHFYDN 60
Db 137 NNVTLSTLKKKNIHSLPKVYFK-YTKLKKIFLOHNCIRHISRAKAFGLCQLIYLNHN 195
Qy 61 PLOVGRSAFOYLPRKLTLSLNGAMDIOEPDLK-----GTTSEILTLTRAGIRLLP 113
Db 196 CTTTLRPGIFKDLQHLTWILD-----DMPITRISQRLFTGLNSLFLSVNNYLEALP 249
Qy 114 SGMCOOLPRILVLELHNOIEL--PSLHRCQKLEIGLOHNRITWEIGADTFQSLSLSQA 171
Db 250 KOMCAMPOLNWDLEGNRIKYLNTSTFLSCDSLTVLEFRNQGIFVEPKTFSSIKNGE 309
Qy 172 LIDSWAITSIHPEAFSTLSKLDLTDNQLTTLPLAGIGLMHLKLGNALSOAFSK 231
Db 310 LDISSVTTELSEHLFKDLKLKLOKLNLSNP-----LMYLKNGPESLKQLOSL 358
Qy 232 DSFPKRLILEVPA-----YQCCPYGMCASFKAQSGQWEADLHLD 273
Db 359 D-----LERIEIPNITRMFOPMKNLSHIYFKNFRYSYA----- 393
Qy 274 EESKSRPLGLIAQAEHNHYQDDLELOLMEHDSKPHPSVQCSPTPGPKPCEYLEFESMGI 333
Db 394 -----PHVRI-CMPLTLDIGSSEDLANNIL 418
Qy 334 RLAWMAVILLVLCNGVLVLTTFAGGPAPLPVYKVVGA--IAGANTLTJISGCLLASV 390

Db 419 RIFWVIAFITCFGN-----LEVIGMRSTIKMENTHMSIKILCA 460
Qy 391 DALT-----FGOFSEYGARWETGLGRANGELAVLQSEASVLLITLAAVOC 436
Db 461 DCLMGVYLFEVGIEDIKRYGQYKAYALLMESVQCRIMGLFAMLSTEVSVLLITLYLLEK 520
Qy 437 SVSV-----SCVRAKSPSLGSRVAGVGLCCLAG-LAAALPLASV-----GEYGASPLCL 487
Db 521 FLVIVFPESNIRPKGRQTSV-----ILICIMAGFLIAVIFPMNNDYGFNGKNGVCF 574
Qy 488 P-YAPPEGOPALGFTVALVM-MNSFCFLVAVAGYIKLYCDLPRGDPEA--VWDC----- 538
Db 575 PLYDQTEDIGSKGYSGLGFLGVNLAFLIIVFSYITMFCISQKTAQOTTEVANCGRREV 634
Qy 539 AMVRHVAWLIFADGLLYCPVAELFSPASMLGLFPV--TPRAVKS-VLLIVLPLPACINPLLY 596
Db 635 AVANRFFEFVSDAICWIPFVV---KILSLRVEIPDITMTSWIVIFPLVNSALNPILY 691
Qy 597 LLENPFRDDLRL 610
Db 692 TLTTFNFFDKLKL 705

RESULT 13
US-09-928-175-21
; Sequence 21, Application US/09928175
; Patent No. US20020123618A1

; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher J.
; APPLICANT: Gong, Jianhua
; APPLICANT: Daugherty, Betsy
; APPLICANT: Rogers, No. US20020123618A1ma
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 00-1229
; CURRENT APPLICATION NUMBER: US/09/928,175
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/224,455
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 21
; LENGTH: 718
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-928-175-21

Query Match 11.0%; Score 423.5; DB 10; Length 718;
Best Local Similarity 23.7%; Pred. No. 2.4e-25;
Matches 154; Conservative 109; Mismatches 262; Indels 125; Gaps 19;

Qy 4 NLEETLDLWYKLOEPPVAI-RTLGRLOELGFHNNNIKAIPKAFMGNPLLOTIHFYDNPI 62
Db 102 NVTLLSLKKKNIHSLPKVYFSRYTELKRIYLOHNCITHTSRRAFLGHNQIYLSHNCI 161
Qy 63 QVGRSAFOYLPRKLTLSL--NGAMDIOEPDLKGTTSLEILTLTRAGIRLLPSGMCOOL 120
Db 162 TSLRPGIFKDLQHLAWLIDDNPTIRISQ-KSPMGLNSLFLSVNGRRLALPBTLLDAQ 220
Qy 121 PRLVLELHNOIEL--PSLHRCQKLEIGLOHNRITWEIGADTFQSLSLSQA 178
Db 221 POLNWDLANNGIKYITNSTFTLDCSLTVLFLPRNQGIFVEPKTFSSIKNGELDLSNM 280
Qy 179 IRSIHPEAFSTLSKLDLTDNQLTTLPLAGIGLMHLKLGNALSOA-----FSKDSF 234
Db 281 TKLVPYHLESDHLQKLNLSNPILYVHKNQSGSLKQLO--SLDERIEIPNISTGMF 337
Qy 235 PKRLILEVPA--AYQCCPYGMCASFKAQSGQWEADLHLDDESSKRPGLIAQAEHNHY 292
Db 338 QPMKNLSHIYLFKFCST----- 356
Qy 293 DQDLELOLMEHDSKPHPSVQCSPTPGPKPCEYLEFESMGI RLAWMAVILLVLCNGVLVLT 352

Db 357 -----VPHVRI-CMPSTDGISSEEDLLANGILRVSVMVIAFTTCVGNFLVI 401
QY 353 LTVFAGGAPLPPVKFYVGAAGANTLTGISGGLASVDALT-----FGQF 398
Db 402 A-----VRSIIKAENTTHAMSIKILCCADCLMGVYLFVSGVDIKYRGQY 446
QY 399 SEYGARWETGLGCRATGFLAVIGSEASVLLTLTAOVCSVS-----SCVRAKSPSLGS 454
Db 447 OKYALLMMESVPCRLGFLATLTSTEVSVLLTLFLEKFLVIVFPFSNLRGKQTAVAL 506
QY 455 VRAGVLCGLAGLAALPLASV-----GEYGASPLCLPYAPPEGQ-ALAGFTVALVM-M 508
Db 507 ASIMWVGE-----LIAAVPTREDYFCGNFYGNKVCPLHYDQADEFGSRGYSIGIFLG 561
QY 509 NSFCFLVAGAYIKLYCDLPRGDFEAV-----WDCAMVRHVAMLIADGLLYCPVAF 561
Db 562 NLAFELIVISYVTMCSIHKTALQTAEVBSHIGKEVAVANREFVIFSDAICWIPVEV 621
QY 562 SFASMLGFLPVTPBAVKS-VLLVVLPLPACINPLLYLFPNHFDDRLRL 610
Db 622 KILSLQV--EIPGTITSMIVFELPVNSALNPILYTLTTSFFKDKLQ 669

RESULT 14

US-09-928-175-20
; Sequence 20, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher J.
; APPLICANT: Gong, Jianhua
; APPLICANT: Daugherty, Betsy
; APPLICANT: Rogers, No. US20020123618A1ma
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
; FILE REFERENCE: 00-1229
; CURRENT APPLICATION NUMBER: US/09/928, 175
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/224,455
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 20
; LENGTH: 737
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-928-175-20

Query Match 11.0%; Score 423.5; DB 10; Length 737;
Best Local Similarity 23.7%; Pred. No. 2,5e-25;
Matches 134; Conservative 109; Mismatches 262; Indels 125; Gaps 19;

QY 4 NLETLNLNKKLOEFPVAI-RTIGRLQELGFHNNNIKAIPKAFMGNDLQTIHFYDNP 62
Db 121 NVTLLSLKXKRIHRLPKVRSRTYELKTYLQHNCTIHISRAFLGLNLILYISHNCT 180
QY 63 QYGRSAFOYLPRKLTLSL--NGAMDIQEFPPDLKGTSLLELTLTRAGIRLLPSGMOOL 120
Db 181 TSLRPGFIFKDLQAWLILDDNPTTRISQ--KSEFGLNSLFLPMVGNRLPALPETLQOM 239
QY 121 PRLVLELSNOIEL--PSLHRCQKLEIGLQHNRIWEIGADFFSOLSSIQALDSLNA 178
Db 240 POLNMYDLANGKIKYITNSFTLTCDSLTVLFLPRNQIGFEPKFTSSKLNGELSSNM 299
QY 179 IRSIHPEASTLSLVKLDLTNOQLTLPLAGLGLMHLKLGMLALSQA-----FSKDSF 234
Db 300 ITFLPVHLSFDHLQKLNSSNPLLYVHKNOFGSLKQLO--SLDERLEIPINISTGMF 356
QY 235 PKIRILEVPR--AYCCPYGMCASFKAASQWAEADLHDEESSKRPGLGLARQAEENY 292
Db 357 QPKKNSHLYLKTFRYCSY----- 375
QY 293 DQDLDELQLEMEDSKHPSPVQCSPTPGPKCEYLFESMGIRLAVMAIVLLSLCNGLYL 352

Db 376 -----VPHVRI-CMPSTDGISSEEDLLANGILRVSVMVIAFTTCVGNFLVI 420
QY 353 LTVFAGGAPLPPVKFYVGAAGANTLTGISGGLASVDALT-----FGQF 398
Db 421 A-----VRSIIKAENTTHAMSIKILCCADCLMGVYLFVSGVDIKYRGQY 465
QY 399 SEYGARWETGLGCRATGFLAVIGSEASVLLTLTAOVCSVS-----SCVRAKSPSLGS 454
Db 466 OKYALLMMESVPCRLGFLATLTSTEVSVLLTLFLEKFLVIVFPFSNLRGKQTAVAL 525
QY 455 VRAGVLCGLAGLAALPLASV-----GEYGASPLCLPYAPPEGQ-ALAGFTVALVM-M 508
Db 526 ASIMWVGE-----LIAAVPTREDYFCGNFYGNKVCPLHYDQADEFGSRGYSIGIFLG 580
QY 509 NSFCFLVAGAYIKLYCDLPRGDFEAV-----WDCAMVRHVAMLIADGLLYCPVAF 561
Db 581 NLAFELIVISYVTMCSIHKTALQTAEVBSHIGKEVAVANREFVIFSDAICWIPVEV 640
QY 562 SFASMLGFLPVTPBAVKS-VLLVVLPLPACINPLLYLFPNHFDDRLRL 610
Db 641 KILSLQV--EIPGTITSMIVFELPVNSALNPILYTLTTSFFKDKLQ 688

RESULT 15

US-09-804-626-6
; Sequence 6, Application US/09804626
; Patent No. US20020128190A1
; GENERAL INFORMATION:
; APPLICANT: Lobel, Leslie
; APPLICANT: Lustbader, Joyce
; TITLE OF INVENTION: EXPRESSION OF PROPERLY FOLDED AND SOLUBLE EXTRACELLULAR DOMAIN
; FILE REFERENCE: 0575/62259/JPM/SHS
; CURRENT APPLICATION NUMBER: US/09/804, 626
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 6
; LENGTH: 699
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-804-626-6

Query Match 10.9%; Score 420; DB 10; Length 699;
Best Local Similarity 24.0%; Pred. No. 4,3e-25;
Matches 148; Conservative 109; Mismatches 286; Indels 74; Gaps 15;

QY 38 IKAIPKAFMG-NPLIQTIFYDNPIQFWGRSAFOYLPRKLTLSLNGAMDIOEFPPDLKGT 96
Db 61 VVIPSQAFRGINEVYIKIELISQIDSLERIEANAFDNL-----LNLS----- 101
QY 97 TLELITLTLRAGIRLLPSGMOOLPRLVLELSHNOIELPS-----LHRCQK 144
Db 102 ---ELITQNTKMLRYTEPGAFLNLPRLKYLSTICNTGIRKPPDYTKVSSSNFLTEICDN 158
QY 145 LE-----EIGLQHNRIWEIGADFFSOLSSIQALDSLNA-IRSIHPEA 186
Db 159 LHITTPGNAFGMNNESTYIKYNGFEEYOSHANG--TTLTSLLEKENVHLEKHNGA 217
QY 187 ESTLSLVKLDLTNOQLTLPLAGLGLMHLKLGMLALSQAQFSKDSFPRKIRILEVYAY 246
Db 218 PRGAGPKTLDISSTRKQALPSYGLESIORLATTSYSLSKLPSREFVNLBATLTTPS 277
QY 247 OCCPYGMCASFKAASQWAEADLHDEESSKRPGLGLARQAEENH--YQDDDELQLEME 304
Db 278 HCCA-----FRNLPYKQNSHSTISENPSQ--CESTVAKVNNKTLYSMALESLSGW 329
QY 305 DSK-----PHPSVQCSPTPGPKCEYLFESMGIRLAVMAIVLLSLCNGLYLTVFAGG 359
Db 330 DYEGFCULP-KTPRCAPREDANPCEDINGDFELARVLMINTALIMGN-MYVLEFLVTS 387
QY 360 PAPLPPVKFVVGALGANTLTGISGGLASVDALTFGQSEYGARWETGLGCRATGFLAY 419

Db 388 RYKLTVPREFLMCNLSFADFCMGLYLLLIASVDSQTKGQYYNHAIDWQTSGCCSTAGFTY 447
QY 420 LGSEASVLLLTIAAVQCSVSVSVRAYGKSPSLGSRAGVLCCLAGLAALPLASGE 479
| | | | : : : : | : | | | |
Db 448 FASELSVYTLVITLERMHTTYAHLDOCKLRLRAHILMLGWLSSLIAMLPVGSN 507
QY 480 YGASPLCLPYAPPEGQPALGFTVALVMNSFCFLVYAGATIKLYCDLPRGDEAY-WDC 538
| | | | : : : : | : | | | |
Db 508 YMKVSIQPPM--DVETTLQVYILTLILNVVAFELICACYIKIYFAVRNPPELMATNKDT 565
QY 539 AMVRHVAVLIPADGILYCPVAFLSFASMLGFPVTPPEAVKSVLLVPLPACLNPLLYLL 598
: : : | | | : : : : : : | : : : | | | :
Db 566 KIAKKAILITFDFTCMAPISFPAISAAKVPPLITYTNSKVLIVLFYPIINSCANPFLYAI 625
QY 599 FNPHERDDLRLRPARG 615
| : | : | : |
Db 626 FTKTFQRDFEFLLSKFG 642

Search completed: November 8, 2002, 19:37:49
Job time : 12.6019 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2002, 19:33:15 ; Search time 15.0533 Seconds
(without alignments)
1438.571 Million cell updates/sec

Title: US-09-851-595-8

Perfect score: 3850
Sequence: 1 GLHNLLETLIDLNKYLQEPFY.....GGLSGGGGFPQSLAFASHV 736

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents-AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfillseq1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
- and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1140	29.6	644	2	US-08-866-757-2
2	1140	29.6	644	4	US-09-153-593-2
3	513.5	13.3	632	4	US-07-757-342D-6
4	498.5	12.9	695	1	US-08-487-886-2
5	498.5	12.9	695	3	US-08-482-855-2
6	498.5	12.9	695	4	US-08-474-986-2
7	468	12.2	764	4	US-07-741-453A-60
8	460	11.9	764	4	US-07-741-453A-54
9	457.5	11.9	764	4	US-07-757-342D-4
10	445	11.6	764	4	US-07-741-453A-61
11	444	11.5	764	4	US-07-757-342D-5
12	443.5	11.5	700	4	US-07-757-342D-3
13	441	11.5	764	4	US-07-741-453A-59
14	413	10.7	674	4	US-07-757-342D-10
15	413	10.7	699	4	US-07-757-342D-2
16	392.5	10.2	792	4	US-07-741-453A-56
17	374	9.7	611	4	US-07-757-342D-8
18	374	9.7	764	4	US-07-757-342D-7
19	373	9.7	764	4	US-07-741-453A-29
20	364	9.5	795	4	US-07-741-453A-55
21	318.5	8.3	420	4	US-08-795-876-33
22	318.5	8.3	423	4	US-08-795-876-38
23	318.5	8.3	436	4	US-08-795-876-2
24	313	8.1	336	1	US-08-118-270-54
25	313	8.1	336	5	PCT-US93-08528-54
26	289.5	7.5	327	1	US-08-118-270-55
27	289.5	7.5	327	5	PCT-US93-08528-55

28	250.5	6.5	332	1	US-08-118-270-53	Sequence 53, Appl
29	250.5	6.5	332	5	PCT-US93-08528-53	Sequence 53, Appl
30	248	6.4	1091	3	US-08-986-485-5	Sequence 5, Appl
31	241.5	6.3	1101	3	US-08-986-485-2	Sequence 2, Appl
32	226	5.9	605	1	US-08-190-802A-49	Sequence 49, Appl
33	226	5.9	605	4	US-09-063-950-5	Sequence 5, Appl
34	226	5.9	605	4	US-08-477-346-49	Sequence 49, Appl
35	226	5.9	605	4	US-08-473-089-49	Sequence 49, Appl
36	226	5.9	605	4	US-08-487-072A-49	Sequence 50, Appl
37	222.5	5.8	603	1	US-08-190-802A-50	Sequence 50, Appl
38	222.5	5.8	603	4	US-08-477-346-50	Sequence 50, Appl
39	222.5	5.8	603	4	US-08-473-089-50	Sequence 50, Appl
40	222.5	5.8	603	4	US-08-487-072A-50	Sequence 50, Appl
41	218	5.7	1523	3	US-09-191-647-2	Sequence 2, Appl
42	218	5.7	1525	4	US-09-540-245A-2	Sequence 2, Appl
43	218	5.7	1525	4	US-09-540-153-2	Sequence 2, Appl
44	207	5.4	353	3	US-08-986-485-6	Sequence 6, Appl
45	205.5	5.3	673	4	US-09-063-950-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-866-757-2
Sequence 2, Application US/08866757
Patent No. 5858716
GENERAL INFORMATION:
APPLICANT: ELISHOURBAGY, NABIL A
APPLICANT: LI, XIAOTONG
APPLICANT: BERGSMAN, DEBK J
TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: RAINIER & PRESTITA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: IBM Compatible
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/866,757
FILING DATE: 30-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: PRESTITA, PAUL F
REGISTRATION NUMBER: 23, 031
REFERENCE/DOCKET NUMBER: GH-70055
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 644 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-866-757-2

Query Match 29.6%; Score 1140; DB 2; Length 644;
Best Local Similarity 40.0%; Pred. No. 7.3e-94;
Matches 247; Conservative 96; Mismatches 214; Indels 60; Gaps 9;

[illegible]

```

? RESULT 2
? US-09-153-593-2
? : Sequence 2, Application US/09153593A
? : Patent No. 6174994
? : GENERAL INFORMATION:
? : APPLICANT: ELSHOUBAGY, NABIL A
? : APPLICANT: LI, XIAOTONG
? : TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
? : FILE REFERENCE: CG-70055-1
? : CURRENT APPLICATION NUMBER: US/09/153,593A
? : CURRENT FILING DATE: 1998-09-15
? : EARLIER APPLICATION NUMBER: 08/866,757
? : EARLIER FILING DATE: 1997-05-30
? : NUMBER OF SEQ ID NOS: 4
? : SOFTWARE: FastSeq for Windows Version 3.0
? : SEQ ID NO 2
? : LENGTH: 664
? : TYPE: PRT
? : ORGANISM: HOMO SAPIENS
? : US-09-153-593-2

```

	Query Match	29.6%	Score 1140;	DB 4;	Length 644;	
	Best Local Similarity	40.08%;	Pred. No.	7.3e-94;		
	Matches 247; Conservative	96;	Mismatches 214;	Indels 60;	Gaps 9	
QY	IQEPDIDKGTSLILITTRAGIRLLPSCMGCOOLPRVLVELSHNOIELPSLHCCKRIE	146	:	: :	:	:
	:::::	:::	:	:	:	:

```

Db      2 VQGFNLTGTVHESLTLTGTKSSIPNNLCOBQKMLPTLDSVYNNIRDLPSFNCHALE 61
QY      147 EIGLOHNRIMWELGADTFEFSQLSLQALDLSWNAIRSIHPEAFSTLSLVKLDITDNOFLT 200
Db      62 EISLOHNOYIOLKEGTFFQGLIRLIRLDSRNLHIEHSARFNTLQPTINLDVFNELTSE 122
QY      207 PLAGLGIMHMKIKGMLALSOAFSKDSPFKLILEVPAVYOCPCYGMCAFFKASGOWEA 266
Db      122 PTBGLNGMLQVLKNGFKLEALAKDPVNLRSLSVPAVYOCCAFWDGDSY----ANLNT 177
QY      267 EDLHLD-----EESKRPLGLLARAQENHYDDDELQLEMDKSRHPSVOCSPPTPEBK 322
Db      178 EDNSLODHSVAQEKGTADAANANTSTLENDEHSQT-----ITHCTSTGAFK 223
QY      323 PCBYLFESMGITLAWAIVLVSLVLCNGILVLTTFVAGPAPLPVYKFWYAGATAGANTLTGI 382
Db      224 PCBYLIGSMIRLTWLFVFLVALFENLLVLTTFASCT--LPSSKLPGLISVSNLFMGI 282
QY      383 SCGLLASVDALTFEGFSEYGARWEHGLGCRAITGFLAYLGSFASVLLTLTAAVQCSVSC 442
Db      283 YTGILTFDAYSWMGFARFEGIMWEGSSCKYTGFLAVFSSSAIFLLMLATVERSLSAKD 342
QY      443 VRAQSPSLGSVRAGVYGCIALAGLAAALPLASVGEYGAAPLCLTPAPBEGOPALGFT 502
Db      343 IKNKNSNHLKQFRVAALLAFLTGATVTCGFLFRHGEYSASPLCLP--PTGEPSTGFT 400
QY      503 VALVMNMFCEFLVYAGATIKLYCDLPRGDEFBAVDMCAMVRHVAMLTFADGLILYCPAFLS 562
Db      401 VTLVLNLSLAFLLAAVITYTKLCMLKEDELSSENSQSMIKHVAMLITFTNCIFCPVAFPS 460
QY      563 FASMLGTPVPEAVKSVLYLVLPCLPCLNIDPLYLLENPHFRD-----LRLRPRAGDSG 618
Db      461 FAPLITATISIEPELMKSVTLFFPLPACPLNVLFVFNPKFKEDMKLLKRRVYKKGGSYS 520
QY      619 PLATAAAGELEK-----SSDQTOALVAFSDV-----LILBSEAGRPBG 659
Db      521 VSISSOGGCTLEODPEYDQCGMYSHLOGLNYVCCDCESLTLKPYSKHLI----- 569
QY      660 LETYGPSPVTLISGQP 676
Db      570 -KSHSCPALVAASCORP 585

```

US-07-757-342D-6
 ; Sequence 6, Application US/07757342D
 ; Patent No. 6218509
 ; GENERAL INFORMATION:
 APPLICANT: IGARASHI, Masao
 ; MINEGISHI, Takashi
 ; NAKAWURA, Kazuo
 TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
 CUSHMAN
 STREET: 130 Water Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: US
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/757,342D
 FILING DATE: 10-Sep-1991
 CLASSIFICATION: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: BUCKLEY, Linda M.
 REGISTRATION NUMBER: 31003

```

1 Sequence 2, Application US/08487886
2 Patent No. 574446
3 GENERAL INFORMATION:
4 APPLICANT: Kelton, Christie Ann
5 APPLICANT: Schweickhardt, Rene Lynn
6 APPLICANT: Cheng, Shirley Vui Yen
7 APPLICANT: Nugent, No. 574444been Patrice
8 TITLE OF INVENTION: Human Follicle Stimulating
9 TITLE OF INVENTION: Hormone Receptor
10 NUMBER OF SEQUENCES: 2
11 CORRESPONDENCE ADDRESS:
12 ADDRESSEE: Stephan P. Williams,
13 ADDRESSEE: Ares-Serono, Inc.
14 STREET: Exchange Place, 37th floor
15 CITY: Boston
16 STATE: MA
17 COUNTRY: USA
18 ZIP: 02109
19 COMPUTER READABLE FORM:
20 MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
21 MEDIUM TYPE: IBM PS/2, model 55 SX
22 COMPUTER: IBM PS/2, model 55 SX
23 OPERATING SYSTEM: MS-DOS version 4.0
24 SOFTWARE: VAX/VMS Mass11 via Kermit to IBM MS-DOS
25 CURRENT APPLICATION DATA:
26 APPLICATION NUMBER: US/08/487,886
27 FILING DATE: 07-JUN-1995
28 CLASSIFICATION: 514
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: 07/670,085
31 FILING DATE: 15-MAR-1991
32 ATTORNEY/AGENT INFORMATION:
33 NAME: Williams, Stephan P.
34 REGISTRATION NUMBER: 28546
35 REFERENCE/DOCKET NUMBER: US/252
36 TELECOMMUNICATION INFORMATION:
37 TELEPHONE: (617) 723-1300
38 TELEFAX: (617) 723-8923
39 INFORMATION FOR SEQ ID NO: 2:
40 SEQUENCE CHARACTERISTICS:
41 LENGTH: 695
42 TYPE: Amino acid
43 TOPOLOGY: Linear
44 MOLECULE TYPE: protein
45 FEATURE:
46 NAME/KEY: signal sequence
47 LOCATION: -17 to -1
48 IDENTIFICATION METHOD: hydrophobic
49 FEATURE:
50 NAME/KEY: putative amino-terminal extracellular domain
51 LOCATION: 1 to 349
52 IDENTIFICATION METHOD: similarity with other
53 IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular
54 IDENTIFICATION METHOD: domains, hydrophilic
55 FEATURE:
56 NAME/KEY: transmembrane domain
57 LOCATION: 350 to 613
58 IDENTIFICATION METHOD: similarity to other G
59 IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains
60 IDENTIFICATION METHOD: protein-coupled receptor
61 FEATURE:
62 NAME/KEY: putative transmembrane region I
63 LOCATION: 350 to 370
64 IDENTIFICATION METHOD: similarity to other G
65 IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
66 IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
67 IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
68 FEATURE:
69 NAME/KEY: putative transmembrane region II
70 LOCATION: 382 to 404
71 IDENTIFICATION METHOD: similarity to other G
72 IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions
73 IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
74 FEATURE:
75 NAME/KEY: putative transmembrane region III
76 LOCATION: 427 to 448

```

```

IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
FEATURE: hydrophobic, about 20-23 amino acids in length
NAME/KEY: putative transmembrane region IV
LOCATION: 469 to 491
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
FEATURE: hydrophobic, about 20-23 amino acids in length
NAME/KEY: putative transmembrane region V
LOCATION: 512 to 533
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
FEATURE: hydrophobic, about 20-23 amino acids in length
NAME/KEY: putative transmembrane region VI
LOCATION: 557 to 580
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
FEATURE: hydrophobic, about 20-23 amino acids in length
NAME/KEY: putative transmembrane region VII
LOCATION: 592 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
FEATURE: hydrophobic, about 20-23 amino acids in length
NAME/KEY: putative carboxy-terminal intracellular
LOCATION: 614 to 678
US-08-487-886-2

```

```

Query Match      12.9%; Score 498.5; DB 1; Length 695;
Best Local Similarity 24.8%; Pred. No. 4,1e-36;
Matches 163; Conservative 116; Mismatches 284; Indels 93; Gaps 18;

```

```

QY 30 ELGFNNNNIAIEKAFMGNPLQTHFYDNP- QEVGSAFOYLPKLTLSNGADIQ 88
DB 50 ELRFVLTAKLVKOKAFSGFGDEKTEISONDVLEADVFSNLPKHRIKRNANL- 108
QY 89 EEPDLKGTSLLETLTRAGIRLLPSGMCQQLPRLVRLSHNOIELSLHCOQLEET 148
DB 109 -----LYINPEAF-QNLPNLOYLLISNTGINKHLDVHKHISLOKV 147
QY 149 --GLOHN-RIMEIGADTESOLS-SLOALDSNNAIRSHPEAF-----STLH 191
DB 148 LLDIDQDNINHIETKNSVGLSFESVITLKNKGIOEIHNCAFNGQDELMLSDNNNE 207
QY 192 SL-----VKLDLTNOQLTLPPLAGLGLMHLKLGKGLALSOAFSKSPFKRLITL 240
DB 208 ELPDVDFHGASGPVLLDISRTIRHSLPSYGLLENKLLARSTYNNLKLPLEKVALMEA 267
QY 241 EVPAYOCCPYGMCASFKAQGWAEEDLH-----LDDESSKRLPLGLARQAE 289
DB 268 SLTYPSHCANF-----ANMRQISELHPICNKSILRQEVDMYQRRGSSLAENE 319
QY 290 NHHYDQDLEQLEKEDSKRHP--SVQSPTPPGPKPEYLEESGIRLAVVAIVLSVLC 347
DB 320 SSSYSGDMTYTEFDYDLCEVVDVYTCSPKDAFNPCEDINGVILVLIWFISILITG 379
QY 348 NGVLVTFVFGAPLPPEVVGALAGANTLIGISGLLAVDALTFGQFSEGGARRET 407
DB 380 N-LIVLVITTSQYKLVLPFLKCNLAFAADLCIGIYLLIASVDIHTKSGYHNVADIMOT 438
QY 408 GIGCRATGFLAVGSAVLLTLAAVQ-----CSVSCVAVRAYSKPSLGVRAGVLS 461
DB 439 GAGCDAAGFFTVFASLSTVTLATLLERMHTTHAMQDC-----KVQLRHAASVVMVG 493
QY 462 CLALGALALPLASGEYGASPLCLPYRPRGQPALEFTVALVMNSFCELVAVGAYI 521
DB 494 WL-FAPAAALFPIFGISIMKYSICLP--DIDPSLSQLYVMSLLVLANVLAVVIGCII 550

```

```

QY 522 KLYCDLPBGDF-EAYWDCAMVHVAWLIFADGLYCPVAFLSFASMLQFLVTPFAVSV 580
DB 551 LTVLRNPNIVSSSDTRIKRNAMLLFTDFLCMAPISFAISASLKVPLITYSKAIL 610
QY 581 LTVLPACNPLLYLLENPFRRDLRLRPARGDSGLAYAAAGELKSSCST 636
DB 611 LVLPHPNSCANPLLYAIFTKNFRDRDFILLKSG-----CYEMQAOIYRTSTST 661

RESULT 5
US-08-482-855-2
; Sequence 2, Application US/08482855
; Patent No. 6121016
; GENERAL INFORMATION:
; APPLICANT: Kelton, Christie Ann
; APPLICANT: Schweickhardt, Rene Lynn
; APPLICANT: Cheng, Shirley Yui Yen
; APPLICANT: Nugent, No. 6121016een Patrice
; TITLE OF INVENTION: Human Follicle Stimulating
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Stephan P. Williams,
; ADDRESSEE: Ares-Serono, Inc.
; STREET: Exchange Place, 37th floor
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
; COMPUTER: IBM PS/2, model 55 SX
; OPERATING SYSTEM: MS-DOS version 4.0
; SOFTWARE: VAX/VMS Massll via Kermit to IBM MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,855
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 07/670,085
; FILING DATE: 15-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Stephan P.
; REGISTRATION NUMBER: 28546
; REFERENCE/DOCKET NUMBER: US/252
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 723-1300
; TELEFAX: (617) 723-8923
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 695
; TYPE: Amino acid
; TOPOLOGY: Linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: signal sequence
; LOCATION: -17 to -1
; IDENTIFICATION METHOD: hydrophobic
; FEATURE:
; NAME/KEY: putative amino-terminal extracellular domain
; LOCATION: 1 to 349
; IDENTIFICATION METHOD: similarity with other
; IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular
; IDENTIFICATION METHOD: domains, hydrophilic
; FEATURE:
; NAME/KEY: transmembrane domain
; LOCATION: 350 to 613
; IDENTIFICATION METHOD: similarity to other G
; IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains
; NAME/KEY: putative transmembrane region I
; LOCATION: 350 to 370
; IDENTIFICATION METHOD: similarity to other G

```



```

IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region II
LOCATION: 382 to 404
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region III
LOCATION: 427 to 448
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region IV
LOCATION: 469 to 491
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region V
LOCATION: 512 to 533
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VI
LOCATION: 557 to 580
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VII
LOCATION: 592 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative carboxy-terminal intracellular
LOCATION: 614 to 678
US-08-482-855-2

Query Match      12.9%; Score 498.5; DB 3; Length 695;
Best Local Similarity 24.8%; Pred. No. 4,1e-36;
Matches 163; Conservative 116; Mismatches 284; Indels 93; Gaps 18;

QY 30 ELGFHNNNIKAIPKAFMGNPLQTIHFYDNP1-QFVGRSAFOYLPKLTLSLNGAMDIO 88
DB 50 ELRFVLTKLRYOKGAFSGFDLEKIEISONDYLEVLEADVFSNLPKLTREIRKANML- 108
QY 89 EFPPLKGTSLLELTFRAGIRLPLSGMCQQLPRLRYVELSHNOIEELPSLRHOCQLEEI 148
DB 109 -----LYINPEAF-QNLPNIOYLLISNTGKHLPDVHKIHSLOKV 147
QY 149 --GLOHN-RIMEIGADFFSOLS-SLOALDSWNAIRSHPEAF-----SLUH 191
DB 148 LLDIODNINIHTIERNFVGLSFESVILMLNKNGIOETHNCAFNSTQIDELNLSNNNIE 207
QY 192 SL-----VKLDLTDNOLTTPLAGLGLMLKLGKMLALSQAQSKDSFPLRL 240
DB 208 ELPLNDVHGASGAYLIDSKTRISHLSFYGLENLKTLKARSRYNKLKLTLEKLAALMEA 267
QY 241 EWPYAVOCCPYGMCASFPGASQGWAEIDLH-----LDDESSKRPGLGLARQAE 289
DB 268 SLTYVSHCCAF-----ANWRQISELHPICNKSILRQEDVYMTQFGRQSSLAEDNE 319
QY 290 NHYDDDLDELQEMDSRHP--SVQCSPTGPRFPCEVLEFSGWGLRVAWMAIVLTVLC 347
DB 320 SSSRSGFDMTYTEFYDLCNEVVDVTCSPKPAFNPCECDEIMKYNILRYLIMTISLAITG 379
QY 348 NGLVLLTYFAGRPAPLPVKFVGAAGANTLTGISGLLASVDALTFGQFSEYGARMET 407

```

```

DB 380 N-IIVLVITTSQYKLTVPRLMCMIAFADLCIGYLLILASVDIHTHSQTHNYAIDMOT 438
QY 408 GIGCATGFLAVIGSEASVILLETAAVO-----CSVSCYRAVAKSPSLGSVRAGVLG 461
DB 439 GAGCAGAGFTVFASELSVYITLITALEMHTITAMOLDG-----KVQLRAASVMVWG 493
QY 462 CIALAGLAALPLASVGEASPLCLPYAPREGOPALGFYALVAMNSFCFLVYAGANI 521
DB 494 WT-FAFAALPLFEGISSYMKVSLCPM--DIDSPLSQLYVMSLLVNLVLAFFVLCGCTI 550
QY 522 KLYCDLPRDGF-EAWWDCAMVHRVAMLEFADGLVCPVAFSLPMSMLGLFPYTPAEVRSV 580
DB 551 HYLTVRNPNIVSSSDTRIARKMAMLTDTFLCMAPISFPALSAKLKPLITVSAKTL 610
QY 581 LLYVLEPLPACLPPLLYLFPNPHRFDLRRLRPAGDSGGLAYAAAGELEKSSCDST 636
DB 611 LVLFHINSCANFLTAFTKNRDFEILSKCG-----CYEMQAOIVKTERVST 661

RESULT 6
US-08-474-986-2
GENERAL INFORMATION:
APPLICANT: Kellon, Christie Ann
Schweickhardt, Rene Lynn
Cheng, Shirley Wui Yen
Nugent, No. 6372711een Patrice
TITLE OF INVENTION: Human Follicle Stimulating
Hormone Receptor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Stephan P. Williams,
Ares-Serono, Inc.
STREET: Exchange Place, 37th floor
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
COMPUTER: IBM PS/2, model 55 SX
OPERATING SYSTEM: MS-DOS version 4.0
SOFTWARE: VAX/VMS Massll via kermit to IBM MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,986
FILING DATE: 07-Jun-1995
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/670,085
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Stephan P.
REGISTRATION NUMBER: 28546
REFERENCE/DOCKET NUMBER: US/252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 723-1300
TELEFAX: (617) 723-8923
LOCATION: 614 to 678
SEQUENCE DESCRIPTION: SEQ ID NO: 2;
US-08-474-986-2

Query Match      12.9%; Score 498.5; DB 4; Length 695;
Best Local Similarity 24.8%; Pred. No. 4,1e-36;
Matches 163; Conservative 116; Mismatches 284; Indels 93; Gaps 18;

QY 30 ELGFHNNNIKAIPKAFMGNPLQTIHFYDNP1-QFVGRSAFOYLPKLTLSLNGAMDIO 88
DB 50 ELRFVLTKLRYOKGAFSGFDLEKIEISONDYLEVLEADVFSNLPKLTREIRKANML- 108
QY 89 EFPPLKGTSLLELTFRAGIRLPLSGMCQQLPRLRYVELSHNOIEELPSLRHOCQLEEI 148
DB 109 -----LYINPEAF-QNLPNIOYLLISNTGKHLPDVHKIHSLOKV 147

```


Sat Nov 9 13:04:31 2002

us-09-851-595-8.ra1

Page 7

APPLICANT: DIMONT, JACQUES
APPLICANT: VASANT, GILBERT
TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR
TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS
TITLE OF INVENTION: AND POLYPEPTIDES, AND APPLICATIONS OF THESE POLYPEPTIDES
TITLE OF INVENTION: 62
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARDY & CUSHMAN
STREET: 1615 L STREET, N.W.
CITY: WASHINGTON, D.C.
COUNTRY: U.S.A.
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/741,453A
FILING DATE: 19911015
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16773
REFERENCE/DOCKET NUMBER: 91913/1107/US/ST
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 822-0944
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 764 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-741-453A-54

Query Match 11.9%; Score 460; DB 4; Length 764;
Best Local Similarity 23.1%; Pred. No. 1.4e-32; Indels 100; Gaps 18;
Matches 167; Conservative 125; Mismatches 331;

QY 17 EPPVAIRLTGR-----QELGFHNNIKAIPEKAFMNPILQTHF-YDNPIDFVGS 95
DB - 36 DFRVTCCKDHRIPLPPSTQTLKFIETQKTPSRASNPINISRIYSIDATLQRESH 125
QY 69 AFQYLPKRLTSLNGAMDIOEF-PD-LKGTTSLLETLITRAGITLLPS-GMCOQLPRLRV 155
DB 96 SFYMLSKMTHTIEIRNTRSLTSDPDLKELPLKFLGIFNGLGVFDPVTVYSGDVEFI 180
QY 126 LELSHN-OIEELPS--LHRCQKLEIGLOHNRIMEIGADTFESQSLSQADLDSMNA-IR 214
DB 156 LEITDNPYMAISIPANAFQGLCNETLTKLYNNFTSIQSHAFNG--TKDAVILNKKKLS 239
QY 181 SIHPAFSTLHS-LVKLIDLTNDOLITPLAGLGIMHLKKGNLALSOAFSKDSEPKLRI 274
DB 215 AIDKDAFEGVYSGPTLLDVSTYVTAALPSKGLHKLILARNWTLLKPLSLISFLHLNR 257
QY 240 LEVPAVYCCPY-----GMCAF----- 334
DB 275 ADLSYPSHCACAFKNOKIRIGLIESLMCNESIRSLRQKSVNLTNGFPDDEYELGDSH 301
QY 258 --FKASQW-----EADLHLDDESSKRPGLLARQAEHNYDQDLDELQ 389
DB 335 AGYADNQGPDQDTSNSHYVFEEDDELIGQGLKLNQDEETLOAFSDSHDYTV----- 389
QY 302 EMEDSKHPVYQSGPRGPKCEYLFEESNGIRLAVAVILSVLCNGVILITVAFAGPA 361
DB 390 ---CGMEDVWCTPKSDENPCEDINGYKFLRIYVWFSLALGNVVLVLLTSYK 445
QY 362 PLPVPKVVGAAGANTLTIGISGLASVDALFFGQFSEYGARWETGLCCRATGELAVIG 421
DB 446 LTVP-RFLMCNLAFADFCMGWTLILASVDLTHSEYNNHAIMDQGTGPGCINAGFTTVA 504

QY 422 SEASVLTLLTAACVSCVSCVRAKSPSLGSRVAGVGLCTALAGLAADPLASVGEYG 481
DB 505 SELSYTTTLVITLERVAITFAMRLDRIRRAHAYALMGWCCLLALPLVGISSTA 564
QY 482 ASPLCIYAPBPQOPALGFTVALVMNSFCFLVAGATIKLCDD-PRGDFE-AWDCAM 540
DB 565 KVICLPM--DTEPLALAYILVLLNVAIFIVCSQYKIIYVANNPPYNGDDFTK 622
QY 541 YRHVAMLIADGLLYCVAFSLFASMLGIFVTPPEAKSVLLVPLPACLNPLAYLLEN 600
DB 623 AKRAVLIETDFWCMAPISFALSALMKPLITVNSKILLVLEFPLNSCANPLVAF 682
QY 601 PHFPDILRLRPAGDSGGLAYVAAAG-----ELEKSCDSTQALVAESVDLIL 649
DB 683 KAFQDVFILLSKGICKROAOAYRQORVSPKNSAGIQLOKVTYRMRQSLPMODEVELL 742
QY 650 EAS 652
DB 743 ENS 745

RESULT 9
US-07-757-342D-4
Sequence 4; Application US/07757342D
Patent No. 6218509
GENERAL INFORMATION:
APPLICANT: IGARASHI, Masao
MINEGISHI, Takashi
NAKAMURA, Kazuo
TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS & CUSHMAN
STREET: 130 Water Street
CITY: Boston
STATE: Massachusetts
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/07/757,342D
FILING DATE: 10-Sep-1991
APPLICATION NUMBER: 31003
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BOCKLEY, Linda M.
REGISTRATION NUMBER: 41226
REFERENCE/DOCKET NUMBER: 41226
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)523-3400
TELEFAX: (617)523-6440
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 696 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-07-757-342D-4

Query Match 11.9%; Score 457.5; DB 4; Length 696;
Best Local Similarity 24.2%; Pred. No. 2e-32; Indels 103; Gaps 16;
Matches 159; Conservative 109; Mismatches 285;

QY 28 LQELGFHNNIKAIPEKAFMNG-NPLQTHFYDNPIDFVGRSAFOYLPKRLTSLNGAMD 101
DB 51 LSRSLITLITVIVISQARGINEVYKTEISQSDLEKTEANADNL-----LNTS----- 101

[illegible]

```

QY 354 TFFAGSPALPEVKKVVGIIAGANTLTITGSCGLASVDLTFQGFSEYGRMTGCGRA 413
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 438 - ILLTSHKLNPRFLMCLMAFADRCMGMYLLILIASVDLTHTSEYINNAIDWOTGGCNT 496
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 414 TGFALVCGSEASVLLLTLLAAVOCVSVSCVARYGKSPSGSVRAGVLAGLALAGLAALP 473
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 497 AGFFVTFVASELSVTLVITTLERWYALTFPAMRIDRKMRLRHACAIWVGWCCFLLALLP 556
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 474 LASVEYAGSPLCLFYAPPEGOPALGFIYVALVMNMSFCFLYYAGAYITLYCDLPNGDEE 533
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 557 LVGISVYAKVSLCLPM - DTEPLALAYIVFLLTLNIAVFVIVCCYVKIYIVRNPQYN 614
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 534 -AVMDCAMVRHAMVILFADGLTYCPVAFISFASMLGLFPVTPEAVKSVLVYLPACLN 592
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 615 PGDKPTFLAKRNAVLIITFDICMAPISFYALSAIILNKPLITYSNSKILLVLFPLMSCAN 674
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 593 PLLVYLFPNHFEDLIRLRPRAGDSGPLAYAAAG-----ELKSSCDSTGALYA 641
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 675 PELVYIFTEKFEDVFIILSKFGICKRQAAVGRQVPPKNSITDIQVKYTHBROGLNH 734
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 642 FSDVDLLEAS 652
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 735 MEDVYELIEKS 745
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
US-07-757-342D-3
Sequence 3, Application US/07757342D
Patent No. 6218509
GENERAL INFORMATION:
APPLICANT: IGARASHI, Masao
MINEGISHI, Takashi
NAKAMURA, Kazuto
TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
CUSHMAN
STREET: 130 Water Street
CITY: Boston
STATE: Massachusetts
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,342D
FILING DATE: 10-Sep-1991
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BUCKLEY, Linda M.
REGISTRATION NUMBER: 31003
REFERENCE//DOCKET NUMBER: 41226
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)523-3400
TELEFAX: (617)523-6440
TELEX: 200291 STRB UR
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 700 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-07-757-342D-3
Query Match 11.5%; Score 443.5; DB: 4; Length 700;
Best Local Similarity 22.9%; Pred. No. 3.7e-31;
Matches 160; Conservative 117; Mismatches 300; Indels 121; Gaps 18;

```

```

QY 38 IKAIEPKAMFG-NPLIOTIHENDNDIOFVGRSAFOYLPRKLTHTSLINGCAMDIQEPPDLKGT 96
Db 65 VAVIYSQAFRGLEWVAKEIETISQSDSLERIEANAFDNL-----LNLS----- 105
QY 97 TSLIETLTTRAGIRLPLSPGMOQOLPRRLVLELSHNOIEPLS-----LHRCOK 144
Db 106 ---ELLIQMTKNLIIYIEPGAFTNIPRLKYLSICNMGIRPLDVTYKISSSEFNLELIDCN 162
QY 145 LE-----EIGLOHNRIMEIGADPFQSLSLQALDLSMNA-INSIHEDA 186
Db 163 LHITTPIGNAFQGMNNSVTLTKYNGEVEVQSHAENG-TTLISLEKEKINYLEKMHSGA 221
QY 187 FSTLHSLVLYLIDNDNOTLTPLLAGIGLMHLKLGMLALSOAFSKDSFPKRLIEVEPRAY 246
Db 222 FGAGRGPSILDISSTKIQALPFRSHGLESIOTLLALSYSKTLPSKEFTSLTAVTLTPYS 281
QY 247 OCCPYGMCASFFKASGQWEADLHLDBESSKRPGLGLARQAEHN--YDODLDELQLEME 304
Db 282 HCCAFENLPK-----KEGNFSEIFENFQKOCSTVRKADNETLYSAIFEENELSGW 333
QY 305 D-----SKPRPSVQCSPTPPGPKPCEYLPEESGIRLAVNAIVLSVLCNGILVLTTFVAG 359
Db 334 DTDYGCSP-KTIQCAPEBDPAFNPCEDIMGYAFRLKYLIMINILATIGN-LTVLFLVLT 391
QY 360 PAPLPVKEFVVGAIAGANTLTIISCGLLASVDALTFQOSEYGARWETGLGCRATGLAV 419
Db 392 RKKLIVPRFLMCLNSPADFQCMGLYLLILASVDSQTKGYYNHAIWDQTSGCCGAFFTV 451
QY 420 LGSEASVLLLTLLAIVQCSVSVSVARAYGKSPSLGSRVAGVGLCTIALAGLAALPLASGE 479
Db 452 FASELSVYTLTVITLEKMTHTTYAVQLODKLRLRHAPIMLGLWFTSLIATMTPLVIGISN 511
QY 480 YGASPCILPEYAPPEGOPALGFTVALVMNNSFCFLVAGAYIKLYCDLPRGDFEAV-MDC 538
Db 512 YKSVISCLPM--DYESTLSQYIISILINVAFAVICACIIRIYFAVQNDELTAPNKDT 569
QY 539 AMVRVAMLLIFADGLYCVAFLSPASMLGLFPVTPPEAKSVLLVLPPLACNPLLYLL 598
Db 570 KIAKKMAILLFTFPTCMAPISFFAISAAFKVPLITVNSKILLIVLFYPVNSCANPEFLVAI 629
QY 599 FNPHEFDLRLRPRRAGSGFLAVAAAGLEKSSCDSTOLVAFSDVDLLIEASAPGRPP 658
Db 630 FYKAFORDELLLSRFG-----CCKRRA-----ELYRRK 658
QY 659 GLETV-----GFP-----SVTLSCQOPGAPR 680
Db 659 EFSAYTSNCKNGFPKASNPQATLKLSTVHCQPIPRP 696

RESULT 13
US-07-741-453A-59
: Sequence 59, Application US/07741453A
: Patent No. 6238597
:
: GENERAL INFORMATION:
: APPLICANT: PARRENTIER, MARC
: APPLICANT: LIBERT, FREDERIC
: APPLICANT: DUMONT, JACQUES
: APPLICANT: VASSART, GILBERT
: TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR
: TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS
: NUMBER OF SEQUENCES: 62
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CUSHMAN DARBY & CUSHMAN
: STREET: 1615 I. STREET, N.W.
: CITY: WASHINGTON, D.C.
: COUNTRY: U.S.A.
: ZIP: 20036
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Disk
: OPERATING SYSTEM: PC-DOS/MS-DOS

```

[illegible]

OY 642 FSDVDILEAS 652
| : : |
Db 735 MEDYELIENS 745

RESULT 14
US-07-757-342D-10
; Sequence 10, Application US/07757342D
; Patent No. 6218509

GENERAL INFORMATION:

APPLICANT: IGARASHI, Masao

MINEGISHI, Takashi

NAKAMURA, Kazuo

TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &

STREET: 130 Water Street

CITY: Boston

STATE: Massachusetts

COUNTRY: US

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/757,342D

FILING DATE: 10-Sep-1991

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: BUCKLEY, Linda M.

REGISTRATION NUMBER: 31003

REFERENCE/DOCKET NUMBER: 41226

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)523-3400

TELEFAX: (617)523-6440

TELEX: 200291 STRE UR

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 674 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 10:

US-07-757-342D-10

Query Match 10.7%; Score 413; DB 4; Length 674;

Best Local Similarity 23.8%; Pred. No. 1.9e-28;

Matches 147; Conservative 109; Mismatches 287; Indels 74; Gaps 15;

OY 38 IKAIEKAMG-NPLIOTIHFYDNPLOFGSAFOYLPRKHTLSLNGAMDIQEPFDLKG 96

Db 36 VAVISQARGLNEVYKIEISQIDSLERIEANAFDNL-----LNLS----- 76

OY 97 TSEILITLRAGIRLPSGMCQOLPRRLVLELSHNOIEELPS-----LHRCQK 144

Db 77 ---ELIQNTKMLRYEPGAFILPGKLYLSICNTGIRKRPDDYTKVSSSSNFLEICDN 133

OY 145 LE-----EIGLOHNRIMEIGADTFSQSSLSQALDSWNA-ITSIHP 186

Db 134 LHITTIIPNAGFOGNNESVTLKLYGNGFEVVOASHFNG-TTLTSLLEKENVHLEKMHNGA 192

OY 187 FSLHSLVLDLTDNOLTLPLAGLGLMLKLGMLATSOAFSKDSFPKLRILEVYAY 246

Db 193 FKGATCPKTLIDISSTYKQALPSYGLSIOQLIATSSYLKLPKRSRETFVNLLENTIYPS 252

OY 247 GCCPYMGCAFSPKASQWAEADLHLDDESSKRPGLGLARQAEHN--YDODLDELQLEME 304

Db 253 HCCA-----FNLPLTKEQNFSSHISENFSKO-CESTVVRKVSNTLYSSMLAESELSGW 304

OY 305 DSK-----PHSVCCSPRPFGPKPCXYLFESMGIRLAWATVILSVLNGVLLTVFPAGG 359

Db 305 DYXGFCILP-KTPRCADPEPDFAFNCEEDIMGYDFLRKLTMLNLIAIMGN-MTVLVLITS 362

OY 360 PAPLPVKFVVGAJAGANTLTIGISGLLASVDALTFQGFSEYGARWETGLGCRATGLAV 419

Db 363 RYKLTVPFRLMCNLSFADFQMGVLLILIASYDSOTKQYNNHAIIDMQSGSGTGFFTV 422

OY 420 LGSFASVLLTLAAVQSVSVSVAVGKSPSLGSRVAGVLCALAGLAAALPLASGE 479

Db 423 FASLSVYTLVTLERWHTLTVAIHLDQKLRHAILIMLGWLFSSLIAMLPVGVSN 482

OY 480 YGASPLCLPYPREGQRALGFYALVMMNSFCLVYAGAVYIKLYCDLPGRGFEAV-MWC 538

Db 483 YMKVSIQFPM--DVETTLISQYIILITLILNVAFPLICACYKITFAVNPDLMAFNKDT 540

OY 539 AMVRHVAMLIPADGLLYCPVAFSLFASMLGFPVPEAVKSVLYLVLPACINPLLYLL 598

Db 541 KIAKKMAILITFTDCMAPISFALSAFVKYPLIVTNSKVLVLFYPIINSCANPELVAI 600

OY 599 FNPFRDDLRLRPRAG 615

Db 601 FTKTFQRDFEFLLSKFG 617

RESULT 15
US-07-757-342D-2
; Sequence 2, Application US/07757342D
; Patent No. 6218509

GENERAL INFORMATION:

APPLICANT: IGARASHI, Masao

MINEGISHI, Takashi

NAKAMURA, Kazuo

TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &

STREET: 130 Water Street

CITY: Boston

STATE: Massachusetts

COUNTRY: US

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/757,342D

FILING DATE: 10-Sep-1991

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: BUCKLEY, Linda M.

REGISTRATION NUMBER: 31003

REFERENCE/DOCKET NUMBER: 41226

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)523-3400

TELEFAX: (617)523-6440

TELEX: 200291 STRE UR

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 699 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-07-757-342D-2

Query Match 10.7%; Score 413; DB 4; Length 699;

Best Local Similarity 23.8%; Pred. No. 2e-28;

Matches 147; Conservative 109; Mismatches 287; Indels 74; Gaps 15;

PT Novel isolated large G-protein coupled receptor 6 polypeptide, useful
 PT for diagnosing and treating weight disorder, metabolic disorder,
 PT central nervous system disorder, endocrine disorder and cardiovascular
 disorder -
 XX
 PS Claim 9; Fig 8; 198pp; English.

CC The invention provides isolated large G-protein coupled receptor 6 (LGR6)
 CC polypeptides and polynucleotides encoding them. The LGR6 polypeptides and
 CC polynucleotides are useful as targets for developing modulating agents
 CC that regulate a variety of cellular processes, e.g. neural and endocrine
 CC processes, as well as thermogenesis. They are useful for developing novel
 CC diagnostic and therapeutic agents for LGR6 associated disorders such as
 CC weight disorders (anorexia, obesity), eating disorders, acquired
 CC immunodeficiency syndrome (AIDS)-related wasting, cancer-related wasting,
 CC metabolic disorders, central nervous system disorders (e.g. Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease, Gilles de la
 CC Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis,
 CC obsessive-compulsive disorder, anxiety, panic disorder, schizophrenia,
 CC endocrine disorders (hypo and hyperthyroidism, dwarfism, gigantism,
 CC acromegaly), and cardiovascular disorders (atherosclerosis, restenosis,
 CC hypertension, myocardial infarction, arrhythmia). The LGR6 polypeptides,
 CC polynucleotides and antibodies are useful in screening assays, detection
 CC assays (e.g., chromosomal mapping, tissue typing, forensic biology),
 CC predictive medicine (e.g. diagnostic assays, prognostic assays,
 CC monitoring clinical trials and pharmacogenomics), and in methods of
 CC treatment (e.g. therapeutic and prophylactic). The present sequence
 CC represents a human LGR6 polypeptide.

XX Sequence 736 AA;

Query Match 100.0%; Score 3850; DB 23; Length 736;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLHNETLDLNNKLOEPPVARTLGRLOELGFHNNNTKAIPKFKMNPLOTTHFND 60
 DB 1 GLHNETLDLNNKLOEPPVARTLGRLOELGFHNNNTKAIPKFKMNPLOTTHFND 60
 QY 61 PLOVGRSAFQYLRKLTSLNGANDIOEPDLKGTSLLEITLTTRAGIRLLPSGCOOL 120
 DB 61 PLOVGRSAFQYLRKLTSLNGANDIOEPDLKGTSLLEITLTTRAGIRLLPSGCOOL 120
 QY 121 PRLVLELSHNOIELPSLRHCQKLEETIGLOHNRIMWEGATFEOSSLOALDSMAIR 180
 DB 121 PRLVLELSHNOIELPSLRHCQKLEETIGLOHNRIMWEGATFEOSSLOALDSMAIR 180
 QY 121 PRLVLELSHNOIELPSLRHCQKLEETIGLOHNRIMWEGATFEOSSLOALDSMAIR 180
 DB 121 PRLVLELSHNOIELPSLRHCQKLEETIGLOHNRIMWEGATFEOSSLOALDSMAIR 180
 QY 181 SHPEASTLSHLYKLLDTNOLTTPLAGIGLMLHLKGNLALSOAFSKDSFPKRLIL 240
 DB 181 SHPEASTLSHLYKLLDTNOLTTPLAGIGLMLHLKGNLALSOAFSKDSFPKRLIL 240
 QY 241 EYVYAVQCCPYGMCASFKAAGGWEADLHLDDESSKRLGLARQAEHHYDDDELQ 300
 DB 241 EYVYAVQCCPYGMCASFKAAGGWEADLHLDDESSKRLGLARQAEHHYDDDELQ 300
 QY 241 EYVYAVQCCPYGMCASFKAAGGWEADLHLDDESSKRLGLARQAEHHYDDDELQ 300
 DB 241 EYVYAVQCCPYGMCASFKAAGGWEADLHLDDESSKRLGLARQAEHHYDDDELQ 300
 QY 301 LEMEDSKPHPSVCCPTPGPFKCEYLFEESWGIRLAWAIVLVLSCNGVLLTVFAGP 360
 DB 301 LEMEDSKPHPSVCCPTPGPFKCEYLFEESWGIRLAWAIVLVLSCNGVLLTVFAGP 360
 QY 361 APPLPKFVAGATAGANTLTGTCGGLASVDALTFCQFSYTGARWETGLCCRAATGLAVL 420
 DB 361 APPLPKFVAGATAGANTLTGTCGGLASVDALTFCQFSYTGARWETGLCCRAATGLAVL 420
 QY 421 GSEAVLLTLTAAVQCSVSCRAKSPSLGSAVAGVLCIALGLAALPLASVGEY 480
 DB 421 GSEAVLLTLTAAVQCSVSCRAKSPSLGSAVAGVLCIALGLAALPLASVGEY 480
 QY 481 GASPLCLPAPREGGPAALGFTVALVMNNSFCFLVAVAGATIKLYCDLPBGDEFAVWDCAM 540
 DB 481 GASPLCLPAPREGGPAALGFTVALVMNNSFCFLVAVAGATIKLYCDLPBGDEFAVWDCAM 540
 QY 541 VRHVAMLIFFADGLLYCPVAFILSFASMLGLFPVTPEAVKSVLLVLPPLPACLNPLLYLLEN 600

DB 541 VRHVAMLIFFADGLLYCPVAFILSFASMLGLFPVTPEAVKSVLLVLPPLPACLNPLLYLLEN 600
 QY 601 PHFRDDLRLRPRAGDSGLPAAVAAAGELKSSCDSTQALVAFSDVDLILEASEAGRPGL 660
 DB 601 PHFRDDLRLRPRAGDSGLPAAVAAAGELKSSCDSTQALVAFSDVDLILEASEAGRPGL 660
 QY 661 EYVGFPSVTLISCOQPAAPRLGSHCVPEEGNHNHGNQPSMDGELLARAGSTPAGGLS 720
 DB 661 EYVGFPSVTLISCOQPAAPRLGSHCVPEEGNHNHGNQPSMDGELLARAGSTPAGGLS 720
 QY 721 GGGGFPSPGLAFASHV 736
 DB 721 GGGGFPSPGLAFASHV 736

RESULT 2
 AAG6141
 ID AAG6141 standard; Protein; 967 AA.

AC AAG6141:
 DT 13-MAR-2002 (first entry)
 DE Human LGR6 polypeptide (clone Fbh150881).

KW Large G-protein coupled receptor 6; LGR6; GPCR; immunomodulator;
 KW anti-HIV; cytostatic; nootropic; neuroprotective; antiparkinsonian;
 KW anticonvulsant; neuroleptic; antidepressant; tranquilizer; antihypertensive;
 KW osteopathic; antiarteriosclerotic; vasotropic; hypotensive; cardiac;
 KW antiarrhythmic; anorectic; gene therapy; human.

OS Homo sapiens.
 PN WO200185768-A2.

15-NOV-2001.

08-MAY-2001; 2001WO-US15002.

08-MAY-2000; 2000US-0566588.

(MILL-) MILLENNIUM PHARM INC.

Gu W;

WPI: 2002-055584/07.

N-PSDB: AA167926, AA167927.

PT Novel isolated large G-protein coupled receptor 6 polypeptide, useful
 PT for diagnosing and treating weight disorder, metabolic disorder,
 PT central nervous system disorder, endocrine disorder and cardiovascular
 disorder -

PS Claim 9; Fig 15; 198pp; English.

CC The invention provides isolated large G-protein coupled receptor 6 (LGR6)
 CC polypeptides and polynucleotides encoding them. The LGR6 polypeptides and
 CC polynucleotides are useful as targets for developing modulating agents
 CC that regulate a variety of cellular processes, e.g. neural and endocrine
 CC processes, as well as thermogenesis. They are useful for developing novel
 CC diagnostic and therapeutic agents for LGR6 associated disorders such as
 CC weight disorders (anorexia, obesity), eating disorders, acquired
 CC immunodeficiency syndrome (AIDS)-related wasting, cancer-related wasting,
 CC metabolic disorders, central nervous system disorders (e.g. Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease, Gilles de la
 CC Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis,
 CC obsessive-compulsive disorder, anxiety, panic disorder, migraine),
 CC endocrine disorders (hypo and hyperthyroidism, dwarfism, gigantism,
 CC acromegaly), and cardiovascular disorders (atherosclerosis, restenosis,
 CC hypertension, myocardial infarction, arrhythmia). The LGR6 polypeptides,
 CC polynucleotides and antibodies are useful in screening assays, detection

assays (e.g., chromosomal mapping, tissue typing, forensic biology), predictive medicine (e.g., diagnostic assays, prognostic assays, monitoring clinical trials and pharmacogenomics), and in methods of treatment (e.g., therapeutic and prophylactic). The present sequence represents a human IgH6 polypeptide.

Sequence 967 AA;

Query Match 100.0%; Score 3850; DB 23; Length 967;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GLHNETLDLVNKLQEPVPAIRTLGRLOELGFHNNIKAIPEKAFKGNPLDITTFYDN 60
DB 232 GLHNETLDLVNKLQEPVPAIRTLGRLOELGFHNNIKAIPEKAFKGNPLDITTFYDN 291
QY 61 PLOFVGRSAFOYLPKRLTLNLGAMDIQEPDLKGTSLTTLTTRAGIRLLPSGMCOOL 120
DB 292 PLOFVGRSAFOYLPKRLTLNLGAMDIQEPDLKGTSLTTLTTRAGIRLLPSGMCOOL 351
QY 121 PRLRLELSHNOIELPLSHCOKLEELGLOHNRIMEIGADTFSSQLQALDLSMAIR 180
DB 352 PRLRLELSHNOIELPLSHCOKLEELGLOHNRIMEIGADTFSSQLQALDLSMAIR 411
QY 181 SIHPEFSTLSLVKLDLTNDLTTLPLAGLGLMHLKGNLALSOAFSKDSPKRLTL 240
DB 412 SIHPEFSTLSLVKLDLTNDLTTLPLAGLGLMHLKGNLALSOAFSKDSPKRLTL 471
QY 241 EYVAYVQCCPYGMCASFRRKASQGMFAEDLHDDSSKRPGLLAROANHYDQDLEIQ 300
DB 472 EYVAYVQCCPYGMCASFRRKASQGMFAEDLHDDSSKRPGLLAROANHYDQDLEIQ 531
QY 301 LEMESKRPSPVQCSPTPQPKFCETLFEFSGIRLVAIVLTVLVCNGLVLTVPAGCP 360
DB 532 LEMESKRPSPVQCSPTPQPKFCETLFEFSGIRLVAIVLTVLVCNGLVLTVPAGCP 591
QY 361 ABLPVKFEVGAIGAGANTLTGISCGLASVDALTGQFSEYGARWETGLGCRATGLAVL 420
DB 592 ABLPVKFEVGAIGAGANTLTGISCGLASVDALTGQFSEYGARWETGLGCRATGLAVL 651
QY 421 GSEASVLTLLTAAVQCSVSCVRAGKSPISGVAVGLGCLALGAAALPLASVGEY 480
DB 652 GSEASVLTLLTAAVQCSVSCVRAGKSPISGVAVGLGCLALGAAALPLASVGEY 711
QY 481 GASPLCLPAPPEGQFALGFTVALVMNNSFCGLVYVAGAYIKLYCDLPRGDFEAVVDCAM 540
DB 712 GASPLCLPAPPEGQFALGFTVALVMNNSFCGLVYVAGAYIKLYCDLPRGDFEAVVDCAM 771
QY 541 VRRHVAWLIFADGLLYCPVAFPLSFASMLGLFPVTPEAVKSVLVVLPPLACNLPLTLLEN 600
DB 772 VRRHVAWLIFADGLLYCPVAFPLSFASMLGLFPVTPEAVKSVLVVLPPLACNLPLTLLEN 831
QY 601 PHRRDLRLRRPAGSGPLAVYAAAGELEKSSCDSTQALVAFSDVLLILEASAGRPGL 660
DB 832 PHRRDLRLRRPAGSGPLAVYAAAGELEKSSCDSTQALVAFSDVLLILEASAGRPGL 891
QY 661 ETVGFPSVTLISQOQAPARLESGHCVEPEGNHFGNPOPSMDGELLIRAGSTPAGGGLS 720
DB 892 ETVGFPSVTLISQOQAPARLESGHCVEPEGNHFGNPOPSMDGELLIRAGSTPAGGGLS 951
QY 721 GGGGFPQSGIAFASHV 736
DB 952 GGGGFPQSGIAFASHV 967

```

RESULT 3
AAE23418
ID AAE23418 standard; Protein; 948 AA.

XX AAE23418;
XX AC
XX DT 27-AUG-2002 (first entry)
XX NO

Human G-protein coupled receptor-10 (GCRG-10) protein.

Human: G-protein coupled receptor-10; GCRG-10; atherosclerosis; cancer; cell proliferative disorder; gastrointestinal; autoimmune; metabolic; neurological; inflammatory; cardiovascular; viral infection; anorexia; cirrhosis; multiple sclerosis; Huntington's disease; gene therapy; Alzheimer's disease; heart failure; hepatitis; cholecystitis; obesity; rheumatoid arthritis; Hashimoto's thyroiditis; angina pectoris; osteoporosis.

Homo sapiens.

Key Location/Qualifiers
Peptide 1..25
/label= signal_peptide
Protein 26..948
/note= "Mature GCRG-10 protein"

Domain 34..65
/note= "Leucine rich repeat N-terminal domain"

Domain 548..571
/label= Transmembrane_domain

Domain 756..783
/label= Transmembrane_domain

MO200226825-A2.

04-APR-2002.

28-SEP-2001; 2001WO-US30661.

29-SEP-2000; 2000US-236546P.

13-OCT-2000; 2000US-240589P.

20-OCT-2000; 2000US-242232P.

03-NOV-2000; 2000US-245855P.

09-NOV-2000; 2000US-247587P.

15-NOV-2000; 2000US-249343P.

(INCY-) INCYTE GENOMICS INC.

Baughn MR, Graul RC, Walla NK, Gandhi AR, Hafalia AJA, Ramkumar J;

Triboley CM, Thornton M, Kallio DA, Yao MG, Elliott VS;

Burford N, Khan FA, Yue H, Lu Y, Arvizu C, Roopa R, Nguyen DB;

Lee EA, Lu DM, Ison CH, Walsh RT, Policky JL;

WPI: 2002-426012/45.

N-PSDB; AAD37674.

Claim 71; Page 126-129; 147p; English.

The invention relates to human G-protein coupled receptor (GCRG 1-16) polypeptides and polynucleotides. Sequences of the invention are useful for the diagnosis, treatment and prevention of cell proliferative (e.g., cancer, cirrhosis), neurological (e.g., multiple sclerosis, Huntington's disease, Alzheimer's disease), cardiovascular (e.g., angina pectoris, heart failure), gastrointestinal (e.g., anorexia, cholecystitis), autoimmune/inflammatory (rheumatoid arthritis, Hashimoto's thyroiditis) and metabolic disorders (e.g., obesity, osteoporosis), viral infections, atherosclerosis and hepatitis. GCRG proteins are useful for identifying compounds that modulate, mimic and block olfactory and taste sensations. They are also useful for identifying GCRG modulators. GCRG DNAs are useful in gene therapy. The present sequence is human GCRG-10 protein.

Query Match 99.6%; Score 3835; DB 23; Length 948;
Best Local Similarity 99.7%; Pred. No. 0;

Matches 734; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GLHNLETLDTLNKLOEPVAIRTLGRLOELGFHNNNNIKAIPEKAFMGNPLLOTTHFYDN 60
 DB 213 GLHNLETLDTLNKLOEPVAIRTLGRLOELGFHNNNNIKAIPEKAFMGNPLLOTTHFYDN 272
 QY 61 PLOFVGRSAFOYLPRKLTHTLSLNGAMDIOEPDLKGTTSLEILTLTRAGIRLLPSGMCQOL 120
 DB 273 TIOFVGRSAFOYLPRKLTHTLSLNGAMDIOEPDLKGTTSLEILTLTRAGIRLLPSGMCQOL 332
 QY 121 PRLRVLELSHNOIEELPSLRQCKLEIGLOHNRIMWEGADTFESQSLQALDLSWNAIR 180
 DB 333 PRLRVLELSHNOIEELPSLRQCKLEIGLOHNRIMWEGADTFESQSLQALDLSWNAIR 392
 QY 181 SIHPEAFSTLHSLVKLDTLTDNQLTTLPLAGIGLMLKLNKLNLSQAFSDSPKRLIL 240
 DB 393 SIHPEAFSTLHSLVKLDTLTDNQLTTLPLAGIGLMLKLNKLNLSQAFSDSPKRLIL 452
 QY 241 EYRYAYOCCPYGMCASFKAAGQWEADLHLDDESSKRPGLLARAENHYDQDLDELQ 300
 DB 453 EYRYAYOCCPYGMCASFKAAGQWEADLHLDDESSKRPGLLARAENHYDQDLDELQ 512
 QY 301 LEMEDSKPHPSVQCSPTPGPFKCEYLFEESWGIRLAWAIVLVLNGVLLTFVAGGP 360
 DB 513 LEMEDSKPHPSVQCSPTPGPFKCEYLFEESWGIRLAWAIVLVLNGVLLTFVAGGP 572
 QY 361 ADLPVKFVVGALAGANTLTGISCGLASVDALTFGQFSEYGARWETGLGCRATGFLAVL 420
 DB 573 ADLPVKFVVGALAGANTLTGISCGLASVDALTFGQFSEYGARWETGLGCRATGFLAVL 632
 QY 421 GSEASVLLTLTAAVQCSVSVSVRAVYKSPSLGSRAGVLCGLAALPLASVGEY 480
 DB 633 GSEASVLLTLTAAVQCSVSVSVRAVYKSPSLGSRAGVLCGLAALPLASVGEY 692
 QY 481 GASPLCLPYAPPEGOPALGFTVALVMNSFCFLVAGATIKLYCDLPRGDFEAVWDCAM 540
 DB 693 GASPLCLPYAPPEGOPALGFTVALVMNSFCFLVAGATIKLYCDLPRGDFEAVWDCAM 752
 QY 541 VRHVMALIFADGLLYCPVAFLSFASMLGLEPVTPEAVKSVLLVVLPLPACLNPLLYLLEN 600
 DB 753 VRHVMALIFADGLLYCPVAFLSFASMLGLEPVTPEAVKSVLLVVLPLPACLNPLLYLLEN 812
 QY 601 PHEFRODLRLRPRAGSGSLAAYAAAGELKSSCDSTQALVAFSDVLLILEASAGRPGL 660
 DB 813 PHEFRODLRLRPRAGSGSLAAYAAAGELKSSCDSTQALVAFSDVLLILEASAGRPGL 872
 QY 661 ETYGPPSVTLISCOQPGARLEGSHCEPEEGNHFGNPOPSMDGELLIRAGSTPAGGGLS 720
 DB 873 ETYGPPSVTLISCOQPGARLEGSHCEPEEGNHFGNPOPSMDGELLIRAGSTPAGGGLS 932
 QY 721 GGGGQPSGLAFASHV 736
 DB 933 GGGGQPSGLAFASHV 948

RESULT 4
 AAU79941 ID AAU79941 standard; Protein: 940 AA.
 AC AAU79941;
 XX 15-JUL-2002 (first entry)
 DE Human orphan GPCR-like protein NOV3.
 XX
 KW Human; NOV3 polypeptide; cardiomyopathy; atherosclerosis; cancer;
 cell signal processing; metabolic pathway modulation; cancerous tissue;
 antibody; diabetes; transgenic animal; orphan GPCR-like protein; NOV3.
 OS Homo sapiens.
 XX
 PN WO200229038-A2.
 XX

PD 11-APR-2002.
 XX
 PF 04-OCT-2001; 2001WO-US31377.
 XX
 PR 04-OCT-2000; 2000US-237862P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Herrmann JL, Rastelli L, Shinkets RA;
 DR WPI; 2002-340104/37.
 DR N-PSDB; AbK49424.
 XX
 PT Novel isolated NOV3 polypeptide, and encoded polynucleotide, useful for
 PT treating cardiomyopathy, atherosclerosis, and cancer
 XX
 PS Claim 1; Page 36; 180pp; English.

The present invention relates to a new NOV3 polypeptide having a
 CC 900 (NOV1), 4349 (NOV2), 940 (NOV3), 798 (NOV4), 865 (NOV5), or
 CC 331 (NOV6) residue amino acid sequence, as given in the specification.
 CC The novel polypeptide, and its encoding polynucleotide, are used to
 CC treat cardiomyopathy, atherosclerosis, cancer or a disease related to
 CC cell signal processing and metabolic pathway modulation, in a human.
 CC detecting the polypeptide or polynucleotide is useful for identifying
 CC cancerous tissue. The antibody can be used to treat diabetes or cancer.
 CC The host cells can be used to produce non-human transgenic animals
 CC useful in drug screening. The present amino acid sequence represents
 CC the human orphan GPCR-like protein NOV3 of the invention.

Sequence 940 AA;
 SQ

Query Match 98.6%; Score 3795; DB 23; Length 940;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 725; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GLHNLETLDTLNKLOEPVAIRTLGRLOELGFHNNNNIKAIPEKAFMGNPLLOTTHFYDN 60
 DB 186 GLHNLETLDTLNKLOEPVAIRTLGRLOELGFHNNNNIKAIPEKAFMGNPLLOTTHFYDN 245
 QY 61 PLOFVGRSAFOYLPRKLTHTLSLNGAMDIOEPDLKGTTSLEILTLTRAGIRLLPSGMCQOL 120
 DB 246 PLOFVGRSAFOYLPRKLTHTLSLNGAMDIOEPDLKGTTSLEILTLTRAGIRLLPSGMCQOL 305
 QY 121 PRLRVLELSHNOIEELPSLRQCKLEIGLOHNRIMWEGADTFESQSLQALDLSWNAIR 180
 DB 306 PRLRVLELSHNOIEELPSLRQCKLEIGLOHNRIMWEGADTFESQSLQALDLSWNAIR 365
 QY 181 SIHPEAFSTLHSLVKLDTLTDNQLTTLPLAGIGLMLKLNKLNLSQAFSDSPKRLIL 240
 DB 366 SIHPEAFSTLHSLVKLDTLTDNQLTTLPLAGIGLMLKLNKLNLSQAFSDSPKRLIL 425
 QY 241 EYRYAYOCCPYGMCASFKAAGQWEADLHLDDESSKRPGLLARAENHYDQDLDELQ 300
 DB 426 EYRYAYOCCPYGMCASFKAAGQWEADLHLDDESSKRPGLLARAENHYDQDLDELQ 485
 QY 301 LEMEDSKPHPSVQCSPTPGPFKCEYLFEESWGIRLAWAIVLVLNGVLLTFVAGGP 360
 DB 486 LEMEDSKPHPSVQCSPTPGPFKCEYLFEESWGIRLAWAIVLVLNGVLLTFVAGGP 545
 QY 361 ADLPVKFVVGALAGANTLTGISCGLASVDALTFGQFSEYGARWETGLGCRATGFLAVL 420
 DB 546 ADLPVKFVVGALAGANTLTGISCGLASVDALTFGQFSEYGARWETGLGCRATGFLAVL 605
 QY 421 GSEASVLLTLTAAVQCSVSVSVRAVYKSPSLGSRAGVLCGLAALPLASVGEY 480
 DB 606 GSEASVLLTLTAAVQCSVSVSVRAVYKSPSLGSRAGVLCGLAALPLASVGEY 665
 QY 481 GASPLCLPYAPPEGOPALGFTVALVMNSFCFLVAGATIKLYCDLPRGDFEAVWDCAM 540
 DB 666 GASPLCLPYAPPEGOPALGFTVALVMNSFCFLVAGATIKLYCDLPRGDFEAVWDCAM 725
 QY 541 VRHVMALIFADGLLYCPVAFLSFASMLGLEPVTPEAVKSVLLVVLPLPACLNPLLYLLEN 600

```

|||||
Db 726 VRHVAWLIFADGLLYCPVAFISFASMLGLFPVTPEAVKSVLLVLPACPACLNPLLYLLEFN 785
QY 601 PHEHDDLRRLPRAGDSGLPVAAGLEKSSCDSTQALVAFSDVLLIEASAGRPPGI 660
Db 786 PHEHDDLRRLPRAGDSGLPVAAGLEKSSCDSTQALVAFSDVLLIEASAGRPPGI 845
QY 661 ETYGFPSTVTLISCOOPGAPRLGSHCVPEEGNHFGNPPQPSMDGELLRAEGSTPAGGSL 720
Db 846 ETYGFPSTVTLISCOOPGAPRLGSHCVPEEGNHFGNPPQPSMDGELLRAEGSTPAGGSL 905
QY 721 GGGGFP 727
Db 906 GGAFFSP 912

RESULT 5
AAV53575
ID AAV53575 standard; Protein; 847 AA.
AC AAV53575;
XX 15-FEB-2000 (first entry)
DE Human gonadotropin receptor partial sequence #5.
XX Human gonadotropin receptor; screening; ligand; biomedical research;
KW biochemical research; drug; hormone; reproductive tissue; infertility;
KW contraception.
XX Homo sapiens.
OS Homo sapiens.
PN EP950711-A2.
XX 20-OCT-1999.
PD 02-FEB-1999; 99EP-0200303.
PE 06-FEB-1998; 98EP-0200357.
PR 27-JUL-1998; 98EP-0202519.
PR 24-SEP-1998; 98EP-0203213.
XX (ALKU) AKZO NOBEL NV.
PA Van Der Spek PJ, Heikoop JC;
PI WPI: 1999-563673/48.
DR N-PSDB; AAZ40461.
XX New 7 transmembrane gonadotropin receptors, useful for screening for
PT hormone analogs and drugs -
PS Claim 1; Page 24-28; 38pp; English.
CC Sequences AAV53571-Y53578 represent fragments of 3 novel human
CC gonadotropin receptors. The novel gonadotropin receptors can be used to
CC screen for ligands of the receptors. This screen may be used in
CC biomedical and biochemical research to develop new drugs targeted to the
CC gonadotropin receptors. For example, hormone analogs which activate or
CC inhibit the function of the gonadotropin receptors or classical
CC gonadotropin receptors may be detected. The altered expression or
CC dysfunction of the gonadotropin receptors causes conditions in
CC reproductive tissues. The ligands can be used for the treatment of
CC infertility or for contraception.
CC
XX Sequence 847 AA:
SO
Query Match 97.8%; Score 3767; DB 20; Length 847;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 720; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 2 LNHLETLIDNINYNLOEPVAVIRTLGRLOELGFHNNNIKAIPEKAFMGNPLLOTIHFYDNP 61
I : : : : :

```

```

Db 94 LPSTQSLDNLANKIQEFPVAVIRTLGRLOELGFHNNNIKAIPEKAFMGNPLLOTIHFYDNP 153
QY 62 IQEFGRSAPFOYLPKHLTLISLNGANDIQEFPDLKGTSTSEILITLRAGIRLLPSGCOQLP 121
Db 154 IQEFGRSAPFOYLPKHLTLISLNGANDIQEFPDLKGTSTSEILITLRAGIRLLPSGCOQLP 213
QY 122 RLRLVLESHNOIEELPSLRQOKLEIEIGLNRRITWEIGADFTSOSSLOALDSMNAIS 181
Db 214 RLRLVLESHNOIEELPSLRQOKLEIEIGLNRRITWEIGADFTSOSSLOALDSMNAIS 273
QY 182 IHPEAFSTLHSLVKLDLTDNOLTTPLLAGLGLMLKLGKLNALSOAFSEKDFPKLRILE 241
Db 274 IHPEAFSTLHSLVKLDLTDNOLTTPLLAGLGLMLKLGKLNALSOAFSEKDFPKLRILE 333
QY 242 VPYAYQCCPYGMCASFRRASQGEAEDLHLDDESSKRPGLIARQAEHHYODLDELQ 301
Db 334 VPYAYQCCPYGMCASFRRASQGEAEDLHLDDESSKRPGLIARQAEHHYODLDELQ 393
QY 302 EMEBSKPHPSVQCSPTPGPFKCEYLFESWGIRLAWAIVLVLVLCNGLVLLTFVAGGPA 361
Db 394 EMEBSKPHPSVQCSPTPGPFKCEYLFESWGIRLAWAIVLVLVLCNGLVLLTFVAGGPA 453
QY 362 PLPPYKFEVGAIAAGANTLTGISCGLIASVDALTFQGFSEYGARWETGLCRATGLAVLG 421
Db 454 PLPPYKFEVGAIAAGANTLTGISCGLIASVDALTFQGFSEYGARWETGLCRATGLAVLG 513
QY 422 SEASVLLTLTAAVQCSVSVCAVATGKSPSLGSVRAGVIGCLALAGLAAALPLASVGEYG 481
Db 514 SEASVLLTLTAAVQCSVSVCAVATGKSPSLGSVRAGVIGCLALAGLAAALPLASVGEYG 573
QY 482 ASPICLPYAPPEGCPAALGFTALVMNNSFCFLVAVGAYIKYICDIPRGDFEAVMDCAMV 541
Db 574 ASPICLPYAPPEGCPAALGFTALVMNNSFCFLVAVGAYIKYICDIPRGDFEAVMDCAMV 633
QY 542 RHVAMLIFADGLLYCPVAFISFASMLGLFPVTPEAVKSVLLVLPACPACLNPLLYLLEFN 601
Db 634 RHVAMLIFADGLLYCPVAFISFASMLGLFPVTPEAVKSVLLVLPACPACLNPLLYLLEFN 693
QY 602 HFRDRLRLRPRAGDSGLPVAAGLEKSSCDSTQALVAFSDVLLIEASAGRPPGLE 661
Db 694 HFRDRLRLRPRAGDSGLPVAAGLEKSSCDSTQALVAFSDVLLIEASAGRPPGLE 753
QY 662 TYGFPSTVTLISCOOPGAPRLGSHCVPEEGNHFGNPPQPSMDGELLRAEGSTPAGGSLG 721
Db 754 TYGFPSTVTLISCOOPGAPRLGSHCVPEEGNHFGNPPQPSMDGELLRAEGSTPAGGSLG 813
QY 722 GGGGFP 727
Db 814 GGAFFSP 819

RESULT 6
AAG66138
ID AAG66138 standard; Protein; 967 AA.
AC AAG66138;
XX 13-MAR-2002 (first entry)
DE Mouse IGR6 polypeptide (clone ftmzb048h10).
XX Large G-protein coupled receptor 6; IGR6; GPCR; immunomodulator;
KW anti-HIV; cytostatic; nootropic; neuroprotective; antiparkinsonian;
KW anticonvulsant; neuroleptic; antidepressant; tranquilizer; antiangine;
KW osteopathic; antiarteriosclerotic; vasotropic; hypotensive; cardiac;
KW antiarrhythmic; anorectic; gene therapy; mouse.
XX
XX Mus musculus.
OS
XX Key 1.23 Location/Qualifiers
FH Peptide /note="signal peptide"
FT Protein 24..967

```


PT central nervous system disorder, endocrine disorder and cardiovascular
PT disorder

PS Example 1; Fig 5; 198pp; English.

XX The invention provides isolated large G-protein coupled receptor 6 (LGR6)
CC polypeptides and polynucleotides encoding them. The LGR6 polypeptides and
CC polynucleotides are useful as targets for developing modulating agents and
CC that regulate a variety of cellular processes, e.g. neural and endocrine
CC processes, as well as thermogenesis. They are useful for developing novel
CC diagnostic and therapeutic agents for LGR6 associated disorders such as
CC weight disorders (anorexia, obesity), eating disorders, acquired
CC immunodeficiency syndrome (AIDS)-related wasting, cancer-related wasting,
CC metabolic disorders, central nervous system disorders (e.g. Alzheimer's
CC disease, Parkinson's disease, Huntington's disease, Gilles de la
CC Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis,
CC epilepsy, Jakob-Creutzfeldt disease, depression, schizophrenia,
CC obsessive-compulsive disorder, anxiety, panic disorder, migraine),
CC endocrine disorders (hypo and hyperthyroidism, dwarfism, gigantism,
CC acromegaly), and cardiovascular disorders (atherosclerosis, restenosis,
CC hypertension, myocardial infarction, arrhythmia). The LGR6 polypeptides,
CC polynucleotides and antibodies are useful in screening assays, detection
CC assays (e.g., chromosomal mapping, tissue typing, forensic biology),
CC predictive medicine (e.g. diagnostic assays, prognostic assays,
CC monitoring clinical trials and pharmacogenomics), and in methods of
CC treatment (e.g. therapeutic and prophylactic). The present sequence
CC represents a human LGR6 polypeptide.

XX Sequence 633 AA;

Query Match 78.3%; Score 3015; DB 23; Length 633;

Best Local Similarity 99.3%; Pred. No. 1.2e-261;
Matches 576; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 157 EIGADTFSSLSIQALDLSNMAIRSIHPEAFSTLHSLVKLDTDNQLTTPPLAGIGLMH 216
DB :: |||||
54 KLEXDFFSOLSIQALDLSNMAIRSIHPEAFSTLHSLVKLDTDNQLTTPPLAGIGLMH 113
QY 217 LKLGMLALSQAASKSPFKRLTEVPYAYAOCCPYGMCASFRRKSGMEADLHDEES 276
DB |||||
114 LKLGMLALSQAASKSPFKRLTEVPYAYAOCCPYGMCASFRRKSGMEADLHDEES 173
QY 277 SKRPLGLARQENHNDQDDELQLEMEDSKPHPSVOCSTPPGFKCEYFEESMGIRLA 336
DB |||||
174 SKRPLGLARQENHNDQDDELQLEMEDSKPHPSVOCSTPPGFKCEYFEESMGIRLA 233
QY 337 WVAIVLLSVLCNGVLVLTFFAGGAPAPLPVKFVVGATAGANTLIGISGILLASVDALTFG 396
DB |||||
234 WVAIVLLSVLCNGVLVLTFFAGGAPAPLPVKFVVGATAGANTLIGISGILLASVDALTFG 293
QY 397 QPSEYGARWETGIGCRATGFLAVLGEASVLLTLAAVOCSSVSCVRANGKSPSLGSVR 456
DB |||||
294 QPSEYGARWETGIGCRATGFLAVLGEASVLLTLAAVOCSSVSCVRANGKSPSLGSVR 353
QY 457 AGVLGCLALAGLAAALPLASVGEYASPLCLPYAPPEGOPAAAGFTVALVMNMFCEFLV 516
DB |||||
354 AGVLGCLALAGLAAALPLASVGEYASPLCLPYAPPEGOPAAAGFTVALVMNMFCEFLV 413
QY 517 AGAYIKLYCDLPGRDEAVWDCAMVRHVAWLIFADGILLYCPVAFSLFASMLGLPVPPEA 576
DB |||||
414 AGAYIKLYCDLPGRDEAVWDCAMVRHVAWLIFADGILLYCPVAFSLFASMLGLPVPPEA 473
QY 577 VASVLLVLPPLPACLNPLLYLFPNPRRDLRLRPARGSGPLATVAAAGLEKSSCDST 636
DB |||||
474 VASVLLVLPPLPACLNPLLYLFPNPRRDLRLRPARGSGPLATVAAAGLEKSSCDST 533
QY 637 QALVAFSDVDLIEASEAGRPGLTEYGPFSVTLISCOOPGAPRLTESHCVEPKNHNGN 696
DB |||||
534 QALVAFSDVDLIEASEAGRPGLTEYGPFSVTLISCOOPGAPRLTESHCVEPKNHNGN 593
QY 697 POPSMDELLLRRAEGSTPAGGGLSGGGFQPSGLAFASHV 736
DB |||||
594 POPSMDELLLRRAEGSTPAGGGLSGGGFQPSGLAFASHV 633

RESULT 8
ABG09314
ID ABG09314 standard; Protein; 691 AA.

AC ABG09314;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #9305.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HSE-) HXSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS73501.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

PS Claim 20; SEQ ID NO 39673; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 691 AA;

Query Match 75.0%; Score 2886.5; DB 22; Length 691;

Best Local Similarity 89.1%; Pred. No. 4.8e-250;
Matches 580; Conservative 10; Mismatches 48; Indels 13; Gaps 9;

QY 85 MDIOEPFDIKGTTSEILTLTRAGIRLPLPSGKQQLPRLVLELSHNQIEPLSLHRCOK 144

DB 1 MDIOEPFDIKGTTSEILTLTRAGIRLPLPSGKQQLPRLVLELSHNQIEPLSLHRCOK 60

QY 145 LEEIGLQHNRIWEIGADTFSSLSIQALDLSNMAIRSIHPEAFSTLHSLVKLDTDNQLT 204

```

|||||
Db 61 LEEIGLQHRRIWEIGADTFESQSLSLQALDLSWNAIRSIHBEAFSTLSHVLKDLTTSQLT 120
QY 205 TLPLAGLGIMLTKGNALNSQAFSKDSPKRLILEVPAYOCSPYGMCSFRRKASQW 264
Db 121 TLPLAGLGIMLTKGNALNSQAFSKDSPKRLILEVPAYOCSPYGMCSFRRKASQW 180
QY 265 EAEDLHLDDESSKRPGLLARQAEENHYDQDLDELQLEMEDSKRHPVQCSPTGPFKPC 324
Db 181 EAEDLHLDDESSKRPGLLARQAEENHYDQDLDELQLEMEDSKRHPVQCSPTGPFKPC 240
QY 325 EYLFESWGIKRLAVALVLLSVLCNGVLTLTVFAGGAPLPVPKVVGAIGANTLTGISC 384
Db 241 EYLFESWGIKRLAVALVLLSVLCNGVLTLTVFAGGAPLPVPKVVGAIGANTLTGISC 300
QY 385 GLIASVDALTFGQFSEYRG-ARHMETGLGCRATGFLAVLGSEASVLLTLTAAVQCSVSCV 443
Db 301 GLIASVDALTFGQFSEYRGARHMETGLGCRATGFLAVLGSEASVLLTLTAAVQCSVSCV 360
QY 444 RAYGKPSLGSVRAGVLCGLALAGLAALPLASVGEYASPLCLPYAPPEGOPALGFTV 503
Db 361 RAYGKPSLGSVRAGVLCGLALAGLAALPLASVGEYASPLCLPYAPPEGOPALGFTV 420
QY 504 ALVMMNSCFLVVAQA---YIKLYCDLPFGDFEAVMDCAMVRHV-AN-LIFADGLTICPV 558
Db 421 ALVMMKLLLFSEWVGPLPYHQMWCYCDLPFGDFEAVMDCAMVRHV-AN-LIFADGLTICPV 480
QY 559 AFLSFASMLGLFPVTPPEAVKSVLLVLPPLACLNPLLYLTFNPHFRDRLRLR-PRAGDS 617
Db 481 AFLSFASMLGLFPVTPPEAVKSVLLVLPPLACLNPLLYLTFNPHFRDRLRLR-PRAGDS 540
QY 618 GPLVAAAGEL--EKSSCDS-TQALVAFSDVDL-ILEASEA-GRPGLETYGFPSVTLI 671
Db 541 GPPXPFLCCGREDGASRASCDSYQALVAVPLSMWISFCEASXSMAGPLELETYGFPSVTLI 600
QY 672 SCQOGGARLBESHCVPEBGNHFGNPPQPSMDGELLRLAESSTPAGGGLSG 722
Db 601 SCQOGGARLBESHCVPEBGNHFGNPPQPSMDGELLRLAESSTPAGGGLSG 651

RESULT 9
AA53574
ID AA53574 standard; Protein; 497 AA.
XX
AC AA53574;
XX
DT 15-FEB-2000 (first entry)
XX
DE Human gonadotropin receptor partial sequence #4.
XX
KW Human; gonadotropin receptor; screening; ligand; biomedical research;
KW biochemical research; drug; hormone; reproductive tissue; infertility;
KW contraception.
XX
OS Homo sapiens.
XX
PN EP950711-A2.
XX
PD 20-OCT-1999.
XX
PF 02-FEB-1999; 99EP-0200303.
XX
PR 06-FEB-1998; 98EP-0200357.
XX
PR 27-JUL-1998; 98EP-0202519.
XX
PR 24-SEP-1998; 98EP-0203213.
XX
PA (ALKU ) AKZO NOBEL NV.
XX
PI Van Der Spek PJ, Heikoop JC;
XX
DR WPI; 1999-563673/48.
XX
DR N-PSDB; AA240460.
XX

```

```

PT New 7 transmembrane gonadotropin receptors, useful for screening for
PT hormone analogs and drugs -
XX
PS Claim 1; Page 18-21; 38pp; English.
XX
CC Sequences AA53574-Y53578 represent fragments of 3 novel human
CC gonadotropin receptors. The novel gonadotropin receptors can be used to
CC screen for ligands of the receptors. This screen may be used in
CC biomedical and biochemical research to develop new drugs targeted to the
CC gonadotropin receptors. For example, hormone analogs which activate or
CC inhibit the function of the gonadotropin receptors or classical
CC gonadotropin receptors may be detected. The altered expression or
CC dysfunction of the gonadotropin receptors causes conditions in
CC reproductive tissues. The ligands can be used for the treatment of
CC infertility or for contraception.
XX
SQ Sequence 497 AA:

```

Query Match 63.6%; Score 2450; DB 20; Length 497;

Best Local Similarity 99.6%; Pred. No. 4,7e-211;

Matches 467; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 259 KASGQEAEDLHLDDESSKRPGLLARQAEENHYDQDLDELQLEMEDSKRHPVQCSPTP 318
Db 1 KASGQEAEDLHLDDESSKRPGLLARQAEENHYDQDLDELQLEMEDSKRHPVQCSPTP 60
QY 319 GPPKPEYLFESWGIKRLAVALVLLSVLCNGVLTLTVFAGGAPLPVPKVVGAIGANT 378
Db 61 GPPKPEYLFESWGIKRLAVALVLLSVLCNGVLTLTVFAGGAPLPVPKVVGAIGANT 120
QY 379 LGTISGGLASVDALTFGQFSEYRGARHMETGLGCRATGFLAVLGSEASVLLTLTAAVQCSV 438
Db 121 LGTISGGLASVDALTFGQFSEYRGARHMETGLGCRATGFLAVLGSEASVLLTLTAAVQCSV 180
QY 439 SVSCVARYGKSPSLGSVRAGVLCGLALAGLAALPLASVGEYASPLCLPYAPPEGOPAA 498
Db 181 SVSCVARYGKSPSLGSVRAGVLCGLALAGLAALPLASVGEYASPLCLPYAPPEGOPAA 240
QY 499 LGFTVALVMMNSCFLVVAQAAYIKLYCDLPFGDFEAVMDCAMVRHVANLIFADGLTICPV 558
Db 241 LGFTVALVMMNSCFLVVAQAAYIKLYCDLPFGDFEAVMDCAMVRHVANLIFADGLTICPV 300
QY 559 AFLSFASMLGLFPVTPPEAVKSVLLVLPPLACLNPLLYLTFNPHFRDRLRLR-PRAGDSG 618
Db 301 AFLSFASMLGLFPVTPPEAVKSVLLVLPPLACLNPLLYLTFNPHFRDRLRLR-PRAGDSG 360
QY 619 PLAYAAAGELKSSCDSTQALVAFSDVDLILEASEAGRPPELETYGFPSVTLISCOQPCA 678
Db 361 PLAYAAAGELKSSCDSTQALVAFSDVDLILEASEAGRPPELETYGFPSVTLISCOQPCA 420
QY 679 PRLGSHCVPEBGNHFGNPPQPSMDGELLRLAESSTPAGGGLSGGCFOP 727
Db 421 PRLGSHCVPEBGNHFGNPPQPSMDGELLRLAESSTPAGGGLSGGCFOP 469

RESULT 10
AAG67556
ID AAG67556 standard; Protein; 395 AA.
XX
AC AAG67556;
XX
DT 26-NOV-2001 (first entry)
XX
DE A human guanosine triphosphate binding protein coupled receptor.
XX
KW Human; guanosine triphosphate binding protein coupled receptor;
KW G protein coupled receptor; brain disease; cancer.
XX
OS Homo sapiens.
XX
PI WO200109323-A1.
XX
DR WO200109323-A1.
XX
DR 08-FEB-2001.
XX

```



```

XX 28-JUL-2000; 2000WO-JP05070.
PF
XX 29-JUL-1999; 99JP-0248036.
PR
XX 27-AUG-1999; 99JP-0300253.
PR
XX 18-OCT-1999; 99US-0159590.
PR
XX 11-JAN-2000; 2000JP-0118776.
PR
XX 17-FEB-2000; 2000US-0183322.
PR
XX 02-MAY-2000; 2000JP-0183767.
PR
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Makamatsu A, Nagai K, Otsuki T, Kishimoto T;
PI Yano K, Kanzaki K, Inoue Y;
XX
XX WPI: 2001-570288/64.
DR
XX N-PSDB; AAX78275.
XX
XX New gene encoding guanosine triphosphate binding protein coupled
PT receptor, and the protein and antibodies to it, useful for diagnosis
PT and treatment of disease such as brain disease -
XX
XX Claim 2; Page 48-51; 63pp; Japanese.
XX
XX The present sequence represents a human guanosine triphosphate binding
CC protein coupled receptor. The guanosine triphosphate binding
CC protein coupled receptor protein is useful in the diagnosis, prediction
CC and treatment of disease associated with disorders of G protein coupled
CC receptor protein, and may be useful in brain disease and cancers.
XX
XX Sequence 395 AA;
SQ
XX
XX Query Match 52.8%; Score 2033; DB 22; Length 395;
XX Best Local Similarity 99.5%; Pred. No. 9,6e-174;
XX Matches 393; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 342 LLSVLCNGLVLTTFVAGPAPLPVYKVVVGAIAAGANTLTGISCGLASVDALTFFGQFSEY 401
DB 1 LLSVLCNGLVLTTFVAGPAPLPVYKVVVGAIAAGANTLTGISCGLASVDALTFFGQFSEY 60
XX
XX 402 GARWEGGCGCRATGFLAVIGSEASVLLTLAAVQCSVSVCAVAKSPSLGVRAGVIG 461
DB 61 GARWEGGCGCRATGFLAVIGSEASVLLTLAAVQCSVSVCAVAKSPSLGVRAGVIG 120
XX
XX 462 CALAGLAAALPLASVGEVGAASPLCLPYAPPEGAPALGFTVALVMNSFCFLVAGAYI 521
DB 121 CALAGLAAALPLASVGEVGAASPLCLPYAPPEGAPALGFTVALVMNSFCFLVAGAYI 180
XX
XX 522 KLYCDLPRGDDEAVWDCAMVRHVAWMLIFADGLLYCPVALTSPASMLGLFPVTPDEAKSVL 581
DB 181 KLYCDLPRGDDEAVWDCAMVRHVAWMLIFADGLLYCPVALTSPASMLGLFPVTPDEAKSVL 240
XX
XX 582 LVLPLPACINPLLYLTFNHFRRDLRLRPRAGDSGLPAAAGELKESKOSTOLVA 641
DB 241 LVLPLPACINPLLYLTFNHFRRDLRLRPRAGDSGLPAAAGELKESKOSTOLVA 300
XX
XX 642 FSDVDLILEASEAGRPGLTETGFPSTVTLISCOOPGAPRLGSHCHVPECGNHFGNPPSM 701
DB 301 FSDVDLILEASEAGRPGLTETGFPSTVTLISCOOPGAPRLGSHCHVPECGNHFGNPPSM 360
XX
XX 702 DGEILLRABGSTPAGGGLGGGGFQSPGLAFASHV 736
DB 361 DGEILLRABGSTPAGGGLGGGGFQSPGLAFASHV 395
XX
XX
XX RESULT 11
XX AAW93889
XX ID AAW93889 standard: Protein; 907 AA.
XX AC AAW93889;
XX DT 25-JUN-1999 (first entry)

```

```

XX XX Human HG38 protein.
DE
XX HG38; human; G-protein coupled glycoprotein hormone receptor; brain;
KW endocrine system; skeletal muscle; spinal cord; placenta; development;
KW receptor activity modulator.
XX
XX Homo sapiens.
XX
XX W09915660-A1.
XX
XX 01-APR-1999.
XX
XX 24-SEP-1998; 98WO-US19979.
XX
XX 24-SEP-1997; 97US-0059863.
XX
XX (MERI) MERCK & CO INC.
XX
XX Bailey WJ, Liu Q, McDonald TP;
PI WPI: 1999-254711/21.
DR N-PSDB; AAX23980.
XX
XX Human G-protein coupled glycoprotein hormone receptor HG38
PT
XX
XX Claim 1a; Fig 2; 74pp; English.
XX
XX This invention describes a novel human G-protein coupled glycoprotein
CC hormone receptor, HG38. Glycoprotein hormone receptors are important in
CC the endocrine system and HG38 may be involved in development and function
CC of the skeletal muscle, spinal cord, placenta and to a lesser extent, the
CC brain. The transgenic animal may be useful for studying tissue and
CC temporal specific expression or activity of the HG38 receptor, as well as
CC for studying the ability of a variety of compounds to act as modulators
CC of HG38 receptor activity.
XX
XX Sequence 907 AA;
SQ
XX
XX Query Match 47.9%; Score 1844; DB 20; Length 907;
XX Best Local Similarity 52.2%; Pred. No. 3.1e-156;
XX Matches 360; Conservative 113; Mismatches 186; Indels 30; Gaps 8;
XX
XX 1 GLNPLETLDINYNKIOEPVAVIRTLGRLOELGFHNNIKAIPEKAFGNPDLQTHFYDN 60
DB 232 GLHSLETLDTNKNLDEFPFAIRTLNMLKELGRSNIRSIPEKAFGNPDLQTHFYDN 291
XX
XX 61 PLOFVGRSAFOVLPKLTHTLSNGAMDIOEFPDLKGTTSLETLVTRAGIRLLPSGMCQOL 120
DB 292 PLOFVGRSAFOVLPKLTHTLSNGASQITEFPDLGTANTLESILTLGAQISSLPQVYNOL 351
XX
XX 121 PRLVLELSHNOIEELPSLRCKOLEETIGLOHNRITWEGADTFEQLSIALDLISNAIR 180
DB 352 PNLQVLDLSNLEEDLPSFVCOXLOKIDLRHNEIYKIVTFQDLISLNSLNAWKIA 411
XX
XX 181 SIHPEARSTLSLVKLDLTDNOLTLPLAGIGLIMHKLKGNLISQAQSKSPKLRIL 240
DB 412 ITHNASTLPSLKLKDLSSMLSPITGLHGLTHKLGNHMLQSLISENPELKV 471
XX
XX 241 EVRYAYOCPEYGMCAEFKASGQV-----EADLHLDEBSSRRPGLARQENHYDQ 294
DB 472 EMPYAYOCCAFVGCENYKISNMWNKGDNSMSMDLH-----KKDGMFOAQDE---R 520
XX
XX 295 DLDLQLEEM-EDSKPHRSVYOCSPRPGRPKPEYLPESGIRLAVATVLSLNGVIL 353
DB 521 DLEDFLLDFEEDLKALHSVOCSPSPGPKPEHLLDGLITIGVTTAVLATLGNALVTS 580
XX
XX 354 TVFAGCAPLPVYKVVVGAIAAGANTLTGISCGLASVDALTFFGQFSEVAGARWEGGLG 413
DB 581 TVFR-SPLYISPIKLLIGVIAAVNMLTGVSASVAVLAGVDAFTFGSFARHGAMWENGVC 639
XX
XX 414 TGFLAVIGSEASVLLTLAAVQCSVSVCAVAKSPSLGVRAGVIGLALAGLAAALP 473
XX
XX

```